

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
31 May 2001 (31.05.2001)

PCT

(10) International Publication Number  
**WO 01/38490 A2**

(51) International Patent Classification<sup>7</sup>:

C12N

(21) International Application Number:

PCT/US00/32403

(22) International Filing Date:

28 November 2000 (28.11.2000)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

60/168,151 29 November 1999 (29.11.1999) US

(71) Applicant: THE TRUSTEES OF COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK [US/US]; West 116th Street and Broadway, New York, NY 10027 (US).

(72) Inventor: DALLA-FAVERA, Riccardo; 445 Riverside Drive, Apt. 122, New York, NY 10025 (US).

(74) Agent: WHITE, John, P.; Cooper & Dunham LLP, 1185 Avenue of the Americas, New York, NY 10036 (US).

(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW.

(84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

— Without international search report and to be republished upon receipt of that report.

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.



**WO 01/38490 A2**

(54) Title: ISOLATION OF FIVE NOVEL GENES CODING FOR NEW Fc RECEPTORS-TYPE MELANOMA INVOLVED IN THE PATHOGENESIS OF LYMPHOMA/MELANOMA

(57) Abstract: This invention provides an isolated nucleic acid molecule which encodes immunoglobulin receptor, Immunoglobulin superfamily Receptor Translocation Associated, IRTA, protein. Provided too, are the IRTA proteins encoded by the isolated nucleic acid molecules, IRTA1, IRTA2, IRTA3, IRTA4 or IRTA5 proteins, having the amino acid sequences set forth in any of Figures 18A, 18B-1-18B-3, 18C-1-18C-2, 18D-1-18D-2 or 18E-1-18E-2. Oligonucleotides of the isolated nucleic acid molecules are provided. Antibodies directed to an epitope of a purified IRTA1, IRTA2, IRTA3, IRTA4 or IRTA5 proteins are also provided, as are pharmaceutical compositions comprising such antibodies or oligonucleotides. Methods for detecting a B cell malignancy in a sample from a subject; diagnosing B cell malignancy in a sample from a subject; detecting human IRTA protein in a sample; and treating a subject having a B cell cancer are also provided.

-1-

**ISOLATION OF FIVE NOVEL GENES  
CODING FOR NEW Fc RECEPTORS-TYPE MELANOMA  
INVOLVED IN THE PATHOGENESIS OF LYMPHOMA/MELANOMA**

- 5       This application claims the priority of copending U.S. Provisional Application Serial No. 60/168,151, filed November 29, 1999, the contents of which are hereby incorporated by reference into the present application.
- 10      The invention disclosed was herein made in the course of work under NCI Grant No. CA 44029 from the National Cancer Institute. Accordingly, the U.S. Government has certain rights in this invention.
- 15      Throughout this application, various references are referred to in parentheses. Disclosures of these publications in their entireties are hereby incorporated by reference into this application to more fully describe the state of the art to which this invention pertains. Full  
20      bibliographic citation for these references may be found at the end of this application, preceding the claims.

**BACKGROUND OF THE INVENTION**

- 25      Abnormalities of chromosome 1q21 are common in B cell malignancies, including B cell lymphoma and myeloma, but the genes targeted by these aberrations are largely unknown. By cloning the breakpoints of a t(1;14)(q21;q32) chromosomal translocation in a myeloma cell line, we have  
30      identified two novel genes, *IRTA1* and *IRTA2*, encoding cell surface receptors with homologies to the Fc and Inhibitory

-2-

Receptor families. Both genes are normally expressed in mature B cells, but with different distributions in peripheral lymphoid organs: IRTA1 is expressed in marginal zone B cells, while IRTA2 is also expressed in germinal center centrocytes and in immunoblasts. As the result of the t(1;14) translocation, the IRTA1 signal peptide is fused to the Immunoglobulin C $\alpha$  domain to produce a chimaeric IRTA1/C $\alpha$  fusion protein. In Multiple Myeloma and Burkitt lymphoma cell lines with 1q21 abnormalities, IRTA2 expression is deregulated. Thus, IRTA1 and IRTA2 are novel immunoreceptors with a potentially important role in B cell development and lymphomagenesis.

B-cell Non-Hodgkin's Lymphoma (B-NHL) and Multiple Myeloma (MM) represent a heterogeneous group of malignancies derived from mature B cells with phenotypes corresponding to pre-Germinal Center (GC) (mantle cell), GC (follicular, diffuse large cell, Burkitt's), or post-GC B cells (MM) (for review, Gaidano and Dalla-Favera, 1997; Kuppers et al., 1999). Insights into the pathogenesis of these malignancies have been gained by the identification of recurrent clonal chromosomal abnormalities characteristic for specific disease subtypes. The common consequence of these translocations is the transcriptional deregulation of protooncogenes by their juxtaposition to heterologous transcriptional regulatory elements located in the partner chromosome (Gaidano and Dalla-Favera, 1997). These heterologous transcriptional regulatory elements can be derived from the Immunoglobulin (IG) locus or from other partner chromosomal loci. Examples include MYC in t(8;14)(q24;q32) in Burkitt's lymphoma (BL) (Dalla-Favera et al., 1982; Taub et al., 1982), the CCND1 gene

-3-

deregulated by the t(11;14) (q13;q32) in mantle cell lymphoma (MCL) (Rosenberg et al., 1991) and multiple myeloma (MM) (Ronchetti et al., 1999), BCL2 involved in the t(14;18) (q32;q21) in follicular lymphoma (FL) (Bakhshi et al., 1985), BCL6 in t(3;14) (q27;q32) in diffuse large B cell lymphoma (DLCL) (Ye et al., 1993), as well as FGFR3 in t(4;14) (p16;q32) (Chesi et al., 1997), MAF in t(14;16) (q32;q23) (Chesi et al., 1998) and MUM1/IRF4 in t(6;14) (p25;q32) (Iida et al., 1997) in multiple myeloma (MM). The identification of these oncogenes has offered valuable insights into the pathogenesis and diagnosis of their corresponding malignancies.

Chromosomal abnormalities involving band 1q21-q23 are among the most frequent genetic lesions in both B-NHL and MM. Among NHL subtypes, translocation breakpoints at 1q21-q23, including translocations and duplications, have been reported, often as the single chromosomal abnormality, in 17-20% of follicular and diffuse large B-cell lymphoma (DLCL), in 39% of marginal-zone B cell lymphoma (Offit et al., 1991; Whang-Peng et al., 1995; Cigudosa et al., 1999) and in 27-38% of Burkitt lymphoma, where they represent the second most common cytogenetic abnormality after translocations involving the MYC proto-oncogene (Berger and Bernheim, 1985; Kornblau et al., 1991). Comparative genome hybridization (CGH) has also identified 1q21-q23 as a recurring site for high-level amplification in 10% of DLCL cases (Rao et al., 1998). In MM, trisomy of the 1q21-q32 region has been reported in 20-31% of cases (Sawyer et al., 1995), amplification of the 1q12-qter region in 80% of cell lines and 40% of primary tumors (Avet-Loiseau et al., 1997), and nonrandom

-4-

unbalanced whole-arm translocations of 1q, associated with the multiduplication of the adjacent 1q21-22 region, were found in 23% of patients with abnormal karyotypes (Sawyer et al., 1998).

5

The high frequency of involvement of 1q21 structural rearrangements in B-cell malignancies suggests that this locus may harbor genes critical to the pathogenesis of these diseases. Cloning of a t(1;14) (q21;q32) in a pre-B 10 cell acute lymphoblastic leukemia cell line previously identified a novel gene, *BCL9* deregulated in this single case (Willis et al., 1998), but not involved in other cases. A recent report characterized the t(1;22) (q22;q11) 15 in a follicular lymphoma (FL) cell line and found that the *FCGR2B* locus, encoding the low affinity IgG Fc receptor *FCGRIIB*, was targeted in this cell line and in two additional FL cases (Callanan et al., 2000). Finally, the *MUC1* locus has been identified in proximity of the breakpoint of a t(1;14) (q21;q32) in NHL (Dyomin et al., 2000; Gilles et al., 2000), and *MUC1* locus rearrangements 20 have been found in 6% of NHL with 1q21 abnormalities (Dyomin et al., 2000). These results highlight the heterogeneity of the 1q21 breakpoints and the need to identify additional candidate oncogenes situated in this 25 locus, since the large majority of these alterations remain unexplained.

The aim of this study was to further explore the architecture of 1q21 chromosomal rearrangements in B cell 30 malignancy. To that end, we have employed a molecular cloning approach of the t(1;14) (q21;q32) present in the myeloma cell line FR4. We have identified two novel genes

-5-

that are differentially targeted by 1q21 abnormalities. These genes code for five novel members of the immunoglobulin receptor family, *IRTA1*, *IRTA2*, *IRTA3*, *IRTA4* and *IRTA5* (Immunoglobulin superfamily Receptor Translocation Associated genes 1, 2, 3, 4, and 5), which may be important for normal lymphocyte function and B cell malignancy.

- 6 -

SUMMARY OF THE INVENTION

This invention provides an isolated nucleic acid molecule which encodes immunoglobulin receptor, Immunoglobulin superfamily Receptor Translocation Associated, IRTA, protein.

This invention provides a method of producing an IRTA polypeptide (protein) which comprises: (a) introducing a vector comprising an isolated nucleic acid which encodes an immunoglobulin receptor, Immunoglobulin superfamily Receptor Translocation Associated, IRTA, protein into a suitable host cell; and (b) culturing the resulting cell so as to produce the polypeptide.

This invention provides an isolated nucleic acid molecule comprising at least 15 contiguous nucleotides capable of specifically hybridizing with a unique sequence included within the sequence of the isolated nucleic acid molecule encoding IRTA protein. In an embodiment, the IRTA protein may be IRTA1, IRTA2, IRTA3, IRTA4 or IRTA5 protein, or fragment(s) thereof, having the amino acid sequence set forth in any of Figures 18A, 18B-1-18B-3, 18C-1-18C-2, 18D-1-18D-2 or 18E-1-18E-2, respectively.

This invention provides a method for detecting a B cell malignancy or a type of B cell malignancy in a sample from a subject wherein the B cell malignancy comprises a 1q21 chromosomal rearrangement which comprises: a) obtaining RNA from the sample from the subject; b) contacting the RNA of step (a) with a nucleic acid molecule of at least 15 contiguous nucleotides capable of specifically

- 7 -

hybridizing with a unique sequence included within the sequence of an isolated RNA encoding human IRTA protein selected from the group consisting of human IRTA1, IRTA2, IRTA3, IRTA4 and IRTA5, under conditions permitting  
5 hybridization of the RNA of step (a) with the nucleic acid molecule capable of specifically hybridizing with a unique sequence included within the sequence of an isolated RNA encoding human IRTA protein, wherein the nucleic acid molecule is labeled with a detectable marker; and c)  
10 detecting any hybridization in step (b), wherein detection of hybridization indicates presence of B cell malignancy or a type of B cell malignancy in the sample.

This invention provides an antisense oligonucleotide  
15 having a sequence capable of specifically hybridizing to an mRNA molecule encoding a human IRTA protein so as to prevent overexpression of the mRNA molecule.

This invention provides a purified IRTA1 protein  
20 comprising the amino acid sequence set forth in Figure 18A (SEQ ID NO:1).

This invention provides a purified IRTA2 protein comprising the amino acid sequence set forth in Figures  
25 18B-1-18B-3 (SEQ ID NO:3).

This invention provides a purified IRTA3 protein comprising the amino acid sequence set forth in Figures 18C-1-18C-2 (SEQ ID NO:5).

-8-

This invention provides a purified IRTA4 protein comprising the amino acid sequence set forth in Figures 18D-1-18D-2 (SEQ ID NO: 7).

- 5 This invention provides a purified IRTA5 protein comprising the amino acid sequence set forth in Figures 18E-1-18E-2 (SEQ ID NO: 9).

This invention provides an antibody/antibodies directed to  
10 an epitope of a purified IRTA1, IRTA2, IRTA3, IRTA4 or  
IRTA5 protein, or fragment(s) thereof, having the amino  
acid sequence set forth in any of Figures 18A, 18B-1-18B-  
3, 18C-1-18C-2, 18D-1-18D-2 or 18E-1-18E-2.

- 15 This invention provides an antibody directed to a purified  
IRTA protein selected from the group consisting of IRTA1,  
IRTA2, IRTA3, IRTA4 and IRTA5.

This invention provides a pharmaceutical composition  
20 comprising an amount of the antibody directed to an IRTA  
protein effective to bind to cancer cells expressing an  
IRTA protein selected from the group consisting of human  
IRTA1, IRTA2, IRTA3, IRTA4 and IRTA5 so as to prevent  
growth of the cancer cells and a pharmaceutically  
25 acceptable carrier.

This invention provides a pharmaceutical composition  
comprising an amount of any of the oligonucleotides of  
nucleic acid molecules encoding IRTA proteins described  
30 herein effective to prevent overexpression of a human IRTA  
protein and a pharmaceutically acceptable carrier capable.

-9-

This invention provides a method of diagnosing B cell malignancy which comprises a 1q21 chromosomal rearrangement in a sample from a subject which comprises:

5 a) obtaining the sample from the subject; b) contacting the sample of step (a) with an antibody directed to a purified IRTA protein capable of specifically binding with a human IRTA protein selected from the group consisting of human IRTA1, IRTA2, IRTA3, IRTA4 and IRTA5 IRTA protein on a cell surface of a cancer cell under conditions

10 permitting binding of the antibody with human IRTA protein on the cell surface of the cancer cell, wherein the antibody is labeled with a detectable marker; and c) detecting any binding in step (b), wherein detection of binding indicates a diagnosis of B cell malignancy in the

15 sample.

This invention provides a method of detecting human IRTA protein in a sample which comprises: a) contacting the sample with any of any of the above-described anti-IRTA antibodies under conditions permitting the formation of a complex between the antibody and the IRTA in the sample; and b) detecting the complex formed in step (a), thereby detecting the presence of human IRTA in the sample.

25 This invention provides a method of treating a subject having a B cell cancer which comprises administering to the subject an amount of anti-IRTA antibody effective to bind to cancer cells expressing an IRTA protein so as to prevent growth of the cancer cells and a pharmaceutically acceptable carrier, thereby treating the subject.

-10-

This invention provides a method of treating a subject having a B cell cancer which comprises administering to the subject an amount of an antisense oligonucleotide having a sequence capable of specifically hybridizing to 5 an mRNA molecule encoding a human ITRA protein so as to prevent overexpression of the human ITRA protein, so as to arrest cell growth or induce cell death of cancer cells expressing ITRA protein(s) and a pharmaceutically acceptable carrier, thereby treating the subject.

10

The invention also provides a pharmaceutical composition comprising either an effective amount of any of the oligonucleotides described herein and a pharmaceutically acceptable carrier.

15

The invention also provides a pharmaceutical composition comprising either an effective amount of an antibody directed against an epitope of any ITRA protein described herein and a pharmaceutically acceptable carrier.

20

-11-

**BRIEF DESCRIPTION OF THE FIGURES**

**Figures 1A-1B.** Molecular cloning of the translocation t(1;14) (q21;q32) in the FR4 multiple myeloma cell line. Fig. 1A) Schematic

5 representation of the  $\lambda$ FR4B-5 and  $\lambda$ FR4S-a clones, representing der(14) and der(1) breakpoints, and of the germline IgH and 1q21 loci. Fig. 1B) Nucleotide sequence of the breakpoint junction and its 10 alignment to the corresponding germline regions of chromosome 14. S $\alpha$ , IgA switch region; LCR: 3' IgH locus control region; B, *BamHI*; H, *HindIII*; X, *XhoI*.

15 **Figures 2A-2B.** Genomic map of the 1q21 locus in the vicinity of the FR4 breakpoint. Fig. 2A) Restriction endonuclease map and schematic

20 representation of genomic clones, i.e. bacteriophages (1), P1 artificial chromosomes (PACs) (2), and yeast artificial chromosome (YAC) (3), spanning the germline 1q21 locus at the FR4 breakpoint region (arrowhead). The name of each clone is placed directly on top of its 25 representation. End fragments derived from the PAC and YAC inserts are depicted as circles, with either an SP6/T7 vector orientation (PAC), or left/right arm vector orientation (YAC). The top panel in Fig. 30 1A depicts the genomic organization of two genes surrounding the FR4 breakpoint. The

-12-

two genes were identified by exon trapping of PAC 49A16. They are closely spaced in the genome, within  $\leq$  30 Kb of each other and are named MUM2 and MUM3 (multiple myeloma-2 and 3). In the scheme of their genomic loci, black boxes indicate coding exons, whereas white and light or medium grey boxes indicate non-coding exons. Connecting introns are lines. MUM3 (left) gives rise to three alternatively spliced mRNAs, all sharing a common 5' untranslated region (UTR) but diverse 3' UTRs (marked by different shades). Numbers underneath the boxes identify the order of exons in the cDNA. Exons less than 100 bp are depicted as thin vertical lines. The position and size of each exon was determined by sequencing of genomic PAC and phage clones and by hybridization of cDNA probes to endonuclease-digested clone DNA. PAC and YAC mapping was performed by partial digestion with rare cutting enzymes followed by Pulse-Field-Gel-Electrophoresis and hybridization to internal and end-derived probes. Dashed lines align regions of overlap. S, SacI; H, HindIII; S, Swal; P<sub>c</sub>, PacI; P, PmeI; Fig. 2B) Genethon genetic linkage map of 1q21 in the region of the MUM2/MUM3 locus. Sequence-tagged sites (STS) are ordered in approximate distance previously determined by Dib, C., et al. (1996) *Nature*, 380:162-

-13-

164. STS WI-5435 (in bold) is contained  
within YAC 23GC4 and PAC 49A16. Parallel  
vertical lines represent interrupted  
segments, whose approximate size is  
5 depicted above in megabases (MB). Sizing  
was estimated by the size of nonchimeric  
YAC contigs between two markers. The BCL9  
gene at the centromere was cloned from a  
10 different t(1;14)(q21;q32) breakpoint by  
Willis T.G. et al., (1998) Blood 91,  
6:1873-1881. The FcGRIIA gene is at the  
1q21-q22 chromosomal band border.

**Figures 3A-3C. MUM2 mRNA structure and expression pattern.**

15 Fig. 3A) Schematic representation of MUM2  
mRNA. Pattern-filled, wide boxes represent  
coding domains and narrow empty boxes  
represent untranslated regions. SP, signal  
peptide; EC, extracellular domain; TM,  
20 transmembrane domain; CYT, cytoplasmic  
domain; A(n), polyA tail. The  
extracellular region is composed of four  
immunoglobulin-like domains as depicted.  
Alternative polyadenylation signals  
25 (arrows) generate three MUM2 mRNA species  
(a, b, c) whose length (in Kb) ranges from  
2.6-3.5. Fig. 3B) Northern blot analysis  
of MUM2 mRNA expresion in human tissues of  
the immune system. The cDNA probe used for  
the analysis is shown as a solid bar  
30 underneath the mRNA scheme in Fig. 3A).  
Each lane contains 2 $\mu$ g mRNA of the

-14-

corresponding tissue. On the right side of  
the blot, the position of RNA molecular  
weight markers is depicted. The position  
of MUM2 and GAPDH mRNA transcripts is shown  
by arrows. (A GAPDH probe was included in  
the hybridization as an internal control -  
0.15 ng labelled +50 ng unlabelled probe-).  
The results of this analysis show weak  
expression of MUM2 in lymph node and  
spleen. MUM2 expression was not detected  
in a variety of other human tissues (data  
not shown). Fig. 3C) Northern blot  
analysis of MUM2 expression in total RNA  
from EREB, a conditional EBV-transformed B  
lymphoblastoid cell line. EREB carries the  
EBV genome with an EBNA2-estrogen receptor  
fusion protein, active only in the presence  
of estrogen. For this experiment, cells  
were grown in the presence of estrogen  
(1 $\mu$ g/ml), followed by estrogen withdrawal  
for the indicated times. Upon estrogen  
withdrawal, EREB cells undergo G0/G1  
arrest, determined by the loss of c-myc  
expression. In Fig. 3C, a Northern blot of  
EREB total RNA (10 $\mu$ g per lane) was  
hybridized with the MUM2 cDNA probe shown  
in Fig. 3A and the GAPDH internal control  
probe, as in Fig. 3B. Arrows indicate the  
position of the corresponding mRNAs on the  
EREB blot. a, band c correspond to the  
MUM2 species in panel Fig 3A. The same  
blot was then stripped and reprobed with a

-15-

c-myc cDNA probe (exon 2) to verify cellular G0/G1 arrest. Quantitation of MUM2 mRNA by the use of a phosphorimager densitometric analysis demonstrates a 10-fold increase in their levels within 48 hrs of estrogen withdrawal, suggesting that MUM2 expression is elevated as the cells enter a resting phase.

10 **Figures 4A-4B.** MUM3 mRNA structure and expression pattern.

Fig. 4A) Schematic representation of MUM3 mRNA Pattern-filled, wide boxes represent coding domains and narrow empty or gray boxes represent untranslated regions. SP, signal peptide; EC, extracellular domain; TM, transmembrane domain; CYT, cytoplasmic domain; A(n), polyA tail. The extracellular region is composed of immunoglobulin-like domains, as depicted. Alternative splicing generates four mRNA species with diverse subcellular localization. MUM3-a and -d proteins are secreted, whereas MUM3-b contains a hydrophobic stretch of amino acids at its C-terminus which may serve as a signal for addition of a glycophosphatidyl-inositol anchor (GPI-anchor), as shown. MUM3-c spans the plasma membrane. Sequence identity among species is indicated by identical filling. Fig. 4B) Northern blot analysis of MUM3 mRNA expression in multiple human tissues (left) and in

-16-

various lymphoid and non-lymphoid cell  
lines (right). The cDNA probe used is  
shown as a solid bar below the cDNA scheme  
in Fig. 4A. Each lane contains 2 $\mu$ g mRNA of  
the corresponding tissue or cell line. The  
position of MUM3 and GAPDH mRNA transcripts  
is shown by arrows. (A GAPDH probe was  
included in the hybridization as an  
internal control as described in Fig. 3)  
a, b, c and d correspond to the MUM3 mRNA  
species shown in Fig. 4A. RD, NC42 and  
CB33, Epstein-Barr virus transformed B  
lymphoblastoid cell lines; EREB,  
conditional EBV-transformed B  
lymphoblastoid cell line; FR4, plasma cell  
line; MOLT4 and HUT78, T cell lines; HL60  
and U937, myelomonocytic cell lines; K562,  
erythroid cell line. The results suggest  
that MUM3 is expressed solely in the immune  
system tissues of bone marrow, lymph and  
spleen and in particular in B cells with a  
lymphoblastoid phenotype.

**Figure 5.** Nucleotide and amino acid sequence of human  
MUM2. The deduced amino acid sequence is  
shown above the nucleotide sequence in one-  
letter code and is numbered on the right,  
with position 1 set to the first codon of  
the signal peptide. The predicted signal  
peptidase site was derived by a computer  
algorithm described in Nielsen et al.,  
Protein Engineering 10, 1-6 (1997) and is

-17-

marked by an arrowhead. The polyadenylation signal AATAAA is underlined. Potential sites for N-glycosylation are also underlined in the amino acid sequence. A hydrophobic stretch of 16 amino acids predicted to span the plasma membrane is doubly underlined. Consensus SH2-binding sites are highlighted by a wavy underline.

10

**Figure 6A.** Nucleotide and amino acid sequence of human MUM3-a. The deduced amino acid sequence is shown above the nucleotide sequence in one-letter code and is numbered on the right, with position 1 set to the first codon of the signal peptide. The predicted site for signal peptidase cleavage was derived as previously stated above and is marked by an arrowhead. The polyadenylation signal ATTAAAA is underlined. Potential sites for N-glycosylation are also underlined in the amino acid sequence. The protein lacks a transmembrane domain and is predicted to be secreted.

25

**Figure 6B.** Nucleotide and amino acid sequence of human MUM3-b. The deduced amino acid sequence is shown above the nucleotide sequence in one-letter code and is numbered on the right, with position 1 set to the first codon of the signal peptide. The predicted site for signal peptidase cleavage was derived as

30

-18-

previously stated above and is marked by an arrowhead. The polyadenylation signal AATAAA is underlined. Potential sites for N-glycosylation are underlined in the amino acid sequence.

5

10

15

20

25

**Figure 6C-1-6C-2.** Nucleotide and amino acid sequence of human MUM3-c. The deduced amino acid sequence is shown above the nucleotide sequence in one-letter code and is numbered on the right, with position 1 set to the first codon of the signal peptide. The predicted site for signal peptidase cleavage was derived as previously stated above and is marked by an arrowhead. The polyadenylation signal AATAAA is underlined. Potential sites for N-glycosylation are underlined in the amino acid sequence. A hydrophobic stretch of 23 amino acids predicted to span the plasma membrane is doubly underlined. Consensus SH2-binding sites are highlighted by a wavy underline.

30

**Figures 7A-7C.** t(1;14) (q21;32) in FR4 generates a MUM2/Ca fusion transcript. Fig. 7A) Schematic representation of the der(14) genomic clone λFR4B-5 and of the germline IgHA1 locus. The FR4 breakpoint is marked by an arrow. Filled and open boxes represent the

-19-

MUM2 and Calpha coding and non-coding exons respectively. The position of the MUM2 exon 1 probe used for Northern blot analysis is shown by a bar. Fig. 7B) 5 Northern blot analysis with a MUM2 exon 1 probe on FR4 and additional cell lines detects an abnormal message of 0.8 Kb, selectively in FR4. Arrowheads point to the location of normal MUM2 message in EREB mRNA. JJN3 and U266, myeloma cell lines; 10 EREB, conditional EBV-transformed B lymphoblastoid cell line. Two µg of polyA+ RNA were loaded per lane. Fig. 7C) Nucleotide and amino acid sequence of the 15 MUM2-Ca fusion cDNA in FR4. The cDNA was amplified by RT-PCR from FR4 total RNA using the primers shown in Fig. 7A, and was subsequently subcloned and sequenced. The deduced amino acid sequence is shown above the nucleotide sequence in one-letter code 20 and is numbered on the right with position 1 set to the first codon of the signal peptide. The predicted site for signal peptidase cleavage was derived as previously stated above and is marked by an 25 arrowhead. The polyadenylation signal AATAAA is underlined. The Calpha transmembrane domain is underlined. The MUM2 portion of the cDNA is shown on italics. H, HindIII; B, BamHI; X, XhoI; 30 S $\alpha$ , IgA switch region; EC, extracellular

-20-

region; TM, transmembrane; CYT, cytoplasmic domain.

Figures 8A-8C. Molecular cloning of the translocation t(1;14) (q21;q32) in the FR4 multiple myeloma cell line. Fig. 8A) Schematic representation of the phage clones representing der(14) and der(1) breakpoints and the germline IgH and 1q21 loci. Chromosome 14 sequences are indicated by a solid black line with black boxes representing Cα1 exons. Chromosome 1 sequences are shown as a grey line. The probes used for chromosomal mapping are indicated below the map. Restriction enzyme codes are: B, BamHI; H, HindIII; X, XbaI; S, SacI; E, EcoRI. For enzymes marked by (\*) only sites delineating the probes are shown. Sa: IgA switch region; LCR: 3'IgH locus control region. Fig. 8B) Nucleotide sequence of the breakpoint junctions and their alignment to the corresponding germline regions of chromosomes 14 and 1. Fig. 8C) Left, fluorescence in situ hybridization (FISH) analysis on human normal metaphase spreads with the PAC clone 49A16 (Fig. 13) spanning the germline 1q21 region at the FR4 breakpoint. Right, DAPI stained image from the same metaphase spread.

-21-

**Figures 9A-9B.** Structure of *IRTA1* and *IRTA2* cDNAs. Figs. 9A,9B) Schematic representation of the full-length *IRTA1* (Fig. 9A) and *IRTA2* (Fig. 9B) cDNAs. Pattern-filled, wide boxes represent coding domains and narrow boxes represent untranslated regions (UTR). The predicted site for signal peptidase cleavage is marked by an arrowhead and was derived according to the SignalIP World Wide Web server at <http://www.cbs.dtu.dk/services/SignalIP>. The transmembrane domain prediction algorithm is described in Tusnady et al, 1998. SP, signal peptide; EC, extracellular domain; Ig, immuno-globulin-type; TM, transmembrane domain; CYT, cytoplasmic domain; A(n), polyA tail; GPI, glycophosphatidyl inositol. In (Fig. 9A), arrows in the 3' UTR indicate different polyadenylation addition sites utilized in the *IRTA1* cDNA. In (Fig. 9B), different 3'UTR regions in *IRTA2* isoforms are differentially shaded. Bars underneath the UTR regions in (Fig. 9A) and (Fig. 9B) identify probes used for Northern blot analysis in Figure 12.

**Figures 10A-10B.** Comparison of the amino acid sequences of *IRTA1* and *IRTA2* with members of the Fc Receptor family Fig. 10A) Multiple sequence alignment of the first two (top) and the third (bottom) extracellular Ig-domains of *IRTA1* and

-22-

IRTA2 to Fc receptor family members. The sequences were compared using the ClustalW program (Thompson et al., 1994). Black-shaded boxes indicate conserved aminoacids among all sequences; dark-grey shaded boxes indicate conserved aminoacids among at least half of the sequences; light-shaded boxes indicate conservative substitutions. Fig. 10B) Alignment of the SH2-binding domains of IRTA1 and IRTA2 with the ITAM and ITIM consensus motifs. Conserved aminoacid positions are in bold. Symbol X represents any aminoacid.

Figures 11A-11B-4. IRTA1 expression pattern. Fig. 11A) Left panel. Northern blot analysis of IRTA1 mRNA expression in tissues of the human immune system. Each lane contains 2mg mRNA. The position of RNA molecular weight markers is depicted on the right side of the blot. The positions of the IRTA1 and GAPDH mRNA transcripts are shown by arrows. (A GAPDH probe was included in the hybridization as an internal control-0.15 ng labelled + 50 ng unlabelled probe-). Right Panel. Northern blot analysis of IRTA1 expression in total RNA from the ER/EB

-23-

cell line (10 mg per lane). For this experiment, cells were grown in the presence of estrogen (1mg/ml), followed by estrogen withdrawal for the indicated times. Arrows indicate the positions of the corresponding mRNAs. a, b and c correspond to the IRTA1 differentially polyadenylated species. The same blot was stripped and reprobed with a MYC cDNA probe (exon 2) to verify cellular G<sub>0</sub>/G<sub>1</sub> arrest. Densitometric analysis of IRTA1 mRNA levels is plotted in the adjacent column graph. The cDNA probe used is shown as a solid bar underneath the IRTA1 mRNA scheme in Figure 9A. Fig. 11B-1-11B-4) *In situ* hybridization analysis of IRTA1 expression in serial sections of human tonsil. 1. Sense IRTA1 probe 2. Antisense IRTA1 probe 3. H&E staining 4. Antisense IRTA1 signal superimposed over an H&E stained section. GC, germinal center, MargZ, marginal zone

**Figure 12A-12B-4.** IRTA2 expression pattern. Fig. 12A) Northern blot analysis of IRTA2 mRNA expression in multiple human tissues (left panel) and in various lymphoid and non-lymphoid cell lines (right panel). Each lane contains 2mg mRNA.

-24-

The positions of the *IRTA2* and *GAPDH* transcripts are shown by arrows. a, b, c and d correspond to the alternatively spliced *IRTA2* mRNA isoforms. RD, NC42 and CB33, Epstein-Barr virus transformed B lymphoblastoid cell lines; EREB, conditional EBV-transformed B lymphoblastoid cell line; FR4, plasma cell line; MOLT4 and HUT78, T cell lines; HL60 and U937, myelomonocytic cell lines; K562, erythroid cell line. The cDNA probe used is shown as a solid bar underneath the *IRTA2* mRNA scheme in Figure 9B. Figs. 12B-1-12B-4) *In situ* hybridization analysis of *IRTA2* mRNA expression in human tonsil. Fig. 12B-1. Sense *IRTA2* cDNA probe, Fig. 12B-2. Antisense *IRTA2* cDNA probe, Fig. 12B-3. H&E staining, Fig. 12B-4. Antisense *IRTA2* cDNA probe signal superimposed over H&E stained section. GC, germinal center, MargZ, marginal zone

Figure 13. Map of the germline 1q21 region spanning the FR4 breakpoint and genomic organization of *IRTA1* and *IRTA2*. Primers used to amplify *IRTA1* exons from spleen cDNA are marked by arrowheads on top panel. Black and light boxes indicate coding and non-coding exons respectively. Arrows indicate

-25-

position of *BCL9*, *MUC1*, *IRTA* family and *FCGRIIB* loci. S, SacI; H, HindIII; S, SwaI; P<sub>c</sub>, PacI; P, PmeI; Mb, Megabases

5      **Figures 14A-14D.**      t(1;14) (q21;q32) in FR4 generates an  
10     *IRTA1/Cα* fusion transcript. Fig. 14A) Schematic representation of the der(14) genomic clone 1FR4B-5 and of the germline *IgCα<sub>1</sub>* locus. The FR4 breakpoint is marked by an arrow. Filled and open boxes represent the *IRTA1* and *Cα<sub>1</sub>* coding and non-coding exons respectively. Fig. 14B) Northern blot analysis with an *IRTA1* exon 1 probe (shown by a bar in Fig. 14A) on FR4 and additional cell lines detects an abnormal message in FR4. Arrowheads point to the location of normal *IRTA1* message in ER/EB mRNA. JJN3 and U266, myeloma cell lines. Two mg of polyA+ RNA loaded per lane. Fig. 14C) Schematic representation of the *IRTA1/Cα* fusion cDNA in FR4. The cDNA was amplified by RT-PCR from FR4 total RNA using the primers shown in (Fig. 14A), and sequenced after subcloning. Fig. 14D) SDS/PAGE analysis of immunoprecipitates obtained from vector control transfected and *IRTA1/Cα* transient expression construct transfected 293-T

15

20

25

30

-26-

cells (lanes 1 & 2), or the following  
cell lines: mIgA positive  
lymphoblastoid cell line-Dakiki (lane  
3), FR4 (lane 4), mIgM positive NHL  
cell line-Ramos (lane 5). H, HindIII;  
B, BamHI; X, XbaI; Sa, IgA switch  
region; EC, extracellular region; TM,  
transmembrane; CYT, cytoplasmic

10      **Figures 15A-15B.**      *IRTA2* expression is deregulated in  
cell lines carrying 1q21  
abnormalities. Figs. 15A, 15B) Northern blot analysis of *IRTA2* mRNA  
expression in Burkitt lymphoma (Fig.  
15A) and Multiple Myeloma (Fig. 15B)  
cell lines. The cDNA probe used is  
the same as in Fig. 12. Each lane  
contains 2mg mRNA. The positions of  
the *IRTA2* and *GAPDH* mRNA transcripts  
are shown by dashes and arrows,  
respectively. The relative levels of  
*IRTA2* mRNA expression in the left  
panel (Fig. 15A) were plotted on the  
right panel (Fig. 15A) after  
densitometric analysis and  
normalization versus the *GAPDH* levels.  
The right panel of (Fig. 15B) is a  
summary of the Northern blot analysis  
results.

30      **Figures 16-1-16-4**      *IRTA1* expression in normal lymphoid  
tissue. Paraffin-embedded sections

-27-

from normal human tonsil were stained  
with the following antibodies: Fig.  
16-1) Negative control; Fig. 16-2)  
anti-CD3 mouse monoclonal to detect T  
5 cells; Fig. 16-3) anti-IRTA1 (mIRTA)  
mouse monoclonal; Fig. 16-4) anti-  
IRTA1 (J92884K) rabbit polyclonal.  
IRTA1 positive cells are located in  
the perifollicular and intraepithelial  
10 region of the tonsil, the equivalent  
of the marginal zone in the spleen.

Figure 17      IRTA1 expression in a stomach Mucosa-  
Associated-Lymphoid Tissue (MALT) B cell  
15 lymphoma. A paraffin-embedded section from  
a stomach MALT B cell lymphoma was stained  
with the anti-IRTA1 (mIRTA) mouse  
monoclonal antibody and counterstained with  
H&E. The majority of MALT lymphomas  
20 analyzed were IRTA1 positive. This  
antibody therefore can be an effective tool  
in the differential diagnosis of MALT  
lymphoma. The mIRTA1 antibody may also be  
proven useful in the therapy of this B cell  
25 tumor, similarly to the use of the anti-  
CD20 antibody (Rituximab) in the therapy of  
relapsed CD20-positive lymphomas (Foon K.,  
Cancer J. 6: p273).

30      Figure 18A.      IRTA1 cDNA and the amino acid sequence of  
the encoded IRTA1 protein.

-28-

- Figures 18B-1-18B-3. IRTA2 cDNA and the amino acid sequence of the encoded IRTA2 protein.
- 5 Figures 18C-1-18C-2. IRTA3 cDNA and the amino acid sequence of the encoded IRTA3 protein.
- 10 Figures 18D-1-18D-2. IRTA4 cDNA and the amino acid sequence of the encoded IRTA4 protein.
- 15 Figures 18E-1-18E-2. IRTA5 cDNA and the amino acid sequence of the encoded IRTA5 protein.
- 20
- 25

-29-

**DETAILED DESCRIPTION OF THE INVENTION**

The following standard abbreviations are used throughout the specification to indicate specific nucleotides: C=cytosine; A=adenosine; T=thymidine and G=guanosine.

5

This invention provides an isolated nucleic acid molecule which encodes immunoglobulin receptor, Immunoglobulin superfamily Receptor Translocation Associated, IRTA, protein.

10

As used herein "Immunoglobulin Receptor Translocation Associated" genes, "IRTA" are nucleic acid molecules which encode novel immunoglobulin superfamily cell surface receptors in B cells which are important in B cell development, and whose abnormal expression, e.g. deregulated expression, perturbs cell surface B cell immunological responses and thus is involved in B cell malignancy, including lymphomagenesis.

20 Nucleic acid molecules encoding proteins designate "MUM-2" and "MUM-3" proteins in the First Series of Experiments are now called "IRTA-1" and "IRTA-2" genes, i.e. nucleic acid molecules which encode IRTA-1 and IRTA-2 proteins respectively. IRTA-3, -4 and -5 proteins are members of  
25 the same the immunoglobulin gene superfamily as are the IRTA-1 and IRTA-2 proteins.

In an embodiment of the above-described isolated nucleic acid molecule which encodes immunoglobulin receptor,  
30 Immunoglobulin superfamily Receptor Translocation Associated, IRTA, protein, the encoded IRTA protein is

-30-

IRTA1 protein comprising the amino acid sequence set forth in Figure 18A (SEQ ID NO:1).

In another embodiment of the above-described isolated  
5 nucleic acid molecule, the encoded IRTA protein is IRTA2  
protein comprising the amino acid sequence set forth in  
Figures 18B-1-18B-3 (SEQ ID NO:3).

In a further embodiment of the above-described isolated  
10 nucleic acid molecule, the encoded IRTA protein is IRTA3  
protein comprising the amino acid sequence set forth in  
Figures 18C-1-18C-2 (SEQ ID NO:5).

In yet another embodiment of the above-described isolated  
15 nucleic acid molecule, the encoded IRTA protein is IRTA4  
protein comprising the amino acid sequence set forth in  
Figures 18D-1-18D-2 (SEQ ID NO: 7).

In a still further embodiment of the above-described  
20 isolated nucleic acid molecule, the encoded IRTA protein  
is IRTA5 protein comprising the amino acid sequence set  
forth in Figures 18E-1-18E-2 (SEQ ID NO: 9).

In another embodiment of any of the above-described  
25 isolated nucleic acid molecules, the nucleic acid molecule  
is DNA. In further embodiments, the DNA is cDNA. In  
additional embodiments, the DNA is genomic DNA. In  
another embodiment, the nucleic acid molecule is an RNA  
molecule. In yet another embodiment, the DNA molecule is  
30 cDNA having the nucleotide sequence set forth in Figure  
18A (SEQ ID NO:2). In another embodiment, the DNA molecule  
is cDNA having the nucleotide sequence set forth in Figure

-31-

18A (SEQ ID NO:4). In a further embodiment, the DNA molecule is cDNA having the nucleotide sequence set forth in Figure 18A (SEQ ID NO:6). In another embodiment, the DNA molecule is cDNA having the nucleotide sequence set forth in Figure 18A (SEQ ID NO:8). In an embodiment, the DNA molecule is cDNA having the nucleotide sequence set forth in Figure 18A (SEQ ID NO:10). In preferred embodiments of the isolated nucleic acid molecule, wherein the nucleic acid molecules encode human IRTA1, IRTA2, IRTA3, IRTA4 or IRTA5 protein. In additional embodiments, the nucleic acid molecules encode mammalian IRTA1 protein. The mammalian IRTA1 protein may be murine IRTA1 protein. In another preferred embodiment, the isolated nucleic acid molecules are operatively linked to a promoter of DNA transcription. In yet another preferred embodiment of the isolated nucleic acid molecule, the promoter comprises a bacterial, yeast, insect, plant or mammalian promoter.

This invention provides a vector comprising any of the above-described isolated nucleic acid molecule encoding IRTA proteins, including but not limited to mammalian IRTA proteins, of which human and murine are preferred. In an embodiment, the vector is a plasmid.

This invention provides a host cell comprising the above-described vector comprising any of the above-described isolated nucleic acid molecule encoding IRTA proteins. Preferably, the isolated nucleic acid molecules in such vectors are operatively linked to a promoter of DNA transcription. In another embodiment of the host cell, the cell is selected from a group consisting of a

-32-

bacterial cell, a plant cell, and insect cell and a mammalian cell.

- This invention provides a method of producing an IRTA polypeptide (protein) which comprises: (a) introducing a vector comprising an isolated nucleic acid which encodes an immunoglobulin receptor, Immunoglobulin superfamily Receptor Translocation Associated, IRTA, protein into a suitable host cell; and (b) culturing the resulting cell so as to produce the polypeptide. In further embodiments, the IRTA protein produced by the above-described method may be recovered and in a still further embodiment, may be purified either wholly or partially. In an embodiment the IRTA protein may be any of IRTA1, IRTA2, IRTA3, IRTA4, or IRTA5 protein. In further embodiments, any of the IRTA proteins may be mammalian proteins. In still further embodiments, the mammalian proteins may be human or mouse IRTA proteins.
- IRTA genes (nucleic acid molecules encoding IRTA proteins IRTA1, IRTA2, IRTA3, IRTA4 and IRTA5) are useful for the production of the IRTA proteins encoded thereby. ITRA proteins are useful for production of antibodies; such antibodies are used as reagents for differential diagnosis of lymphoma subtypes in hematopathology. Antibodies directed against IRTA proteins and which bind specifically to IRTA proteins also have therapeutic uses, i.e. to specifically target tumor cells, which may be used and administered similarly to "Rituximab" (an anti-CD20 antibody), which is an antibody approved by the FDA for therapy of relapsed CD20-positive lymphomas (Foon K., Cancer J. 6(5):273). Anti-IRTA1, anti-IRTA2, anti-IRTA3,

-33-

anti-IRTA4 and anti-IRTA5 antibodies are also useful markers for isolation of specific subsets of B cells in researchstudies of normal and tumor B cell biology. Moreover, Anti-IRTA1, anti-IRTA2, anti-IRTA3, anti-IRTA4  
5 and anti-IRTA5 antibodies are useful research reagents to experimentally study the biology of signaling in normal and tumor B cells.

Methods of introducing nucleic acid molecules into cells  
10 are well known to those of skill in the art. Such methods include, for example, the use of viral vectors and calcium phosphate co-precipitation. Accordingly, nucleic acid molecules encoding IRTA proteins IRTA1, IRTA2, IRTA3, IRTA4 and IRTA5 may be introduced into cells for the  
15 production of these IRTA proteins.

Numerous vectors for expressing the inventive proteins IRTA1, IRTA2, IRTA3, IRTA4, and IRTA5, may be employed. Such vectors, including plasmid vectors, cosmid vectors,  
20 bacteriophage vectors and other viruses, are well known in the art. For example, one class of vectors utilizes DNA elements which are derived from animal viruses such as bovine papilloma virus, polyoma virus, adenovirus, vaccinia virus, baculovirus, retroviruses (RSV, MMTV or  
25 MoMLV), Semliki Forest virus or SV40 virus. Additionally, cells which have stably integrated the DNA into their chromosomes may be selected by introducing one or more markers which allow for the selection of transfected host cells. The markers may provide, for example, prototrophy  
30 to an auxotrophic host, biocide resistance or resistance to heavy metals such as copper. The selectable marker gene can be either directly linked to the DNA sequences to

-34-

be expressed, or introduced into the same cell by cotransformation.

Regulatory elements required for expression include  
5 promoter sequences to bind RNA polymerase and transcription initiation sequences for ribosome binding. Additional elements may also be needed for optimal synthesis of mRNA. These additional elements may include splice signals, as well as enhancers and termination  
10 signals. For example, a bacterial expression vector includes a promoter such as the lac promoter and for transcription initiation the Shine-Dalgarno sequence and the start codon AUG. Similarly, a eukaryotic expression vector includes a heterologous or homologous promoter for  
15 RNA polymerase II, a downstream polyadenylation signal, the start codon AUG, and a termination codon for detachment of the ribosome. Such vectors may be obtained commercially or assembled from the sequences described by methods well known in the art, for example the methods  
20 described above for constructing vectors in general.

These vectors may be introduced into a suitable host cell to form a host vector system for producing the inventive proteins. Methods of making host vector systems are well  
25 known to those skilled in the art.

Suitable host cells include, but are not limited to, bacterial cells (including gram positive cells), yeast cells, fungal cells, insect cells and animal cells.  
30 Suitable animal cells include, but are not limited to HeLa cells, Cos cells, CV1 cells and various primary mammalian cells. Numerous mammalian cells may be used as hosts,

-35-

including, but not limited to, the mouse fibroblast cell NIH-3T3 cells, CHO cells, HeLa cells, Ltk<sup>-</sup> cells and COS cells. Mammalian cells may be transfected by methods well known in the art such as calcium phosphate precipitation,  
5 electroporation and microinjection.

This invention provides an isolated nucleic acid molecule comprising at least 15 contiguous nucleotides capable of specifically hybridizing with a unique sequence included  
10 within the sequence of the isolated nucleic acid molecule encoding IRTA protein. In an embodiment, the IRTA protein may be IRTA1, IRTA2, IRTA3, IRTA4 or IRTA5 protein, or fragment(s) thereof, having the amino acid sequence set forth in any of Figures 18A, 18B-1-18B-3, 18C-1-18C-2,  
15 18D-1-18D-2 or 18E-1-18E-2, respectively. In other embodiments, the isolated nucleic acid molecules are labeled with a detectable marker. In still other embodiments of the isolated nucleic acid molecules, the detectable marker is selected from the group consisting of  
20 a radioactive isotope, enzyme, dye, biotin, a fluorescent label or a chemiluminescent label.

This invention provides a method for detecting a B cell malignancy or a type of B cell malignancy in a sample from  
25 a subject wherein the B cell malignancy comprises a 1q21 chromosomal rearrangement which comprises: a) obtaining RNA from the sample from the subject; b) contacting the RNA of step (a) with a nucleic acid molecule of at least  
15 contiguous nucleotides capable of specifically  
30 hybridizing with a unique sequence included within the sequence of an isolated RNA encoding human IRTA protein selected from the group consisting of human IRTA1, IRTA2,

-36-

IRTA3, IRTA4 and IRTA5, under conditions permitting hybridization of the RNA of step (a) with the nucleic acid molecule capable of specifically hybridizing with a unique sequence included within the sequence of an isolated RNA 5 encoding human IRTA protein, wherein the nucleic acid molecule is labeled with a detectable marker; and c) detecting any hybridization in step (b), wherein detection of hybridization indicates presence of B cell malignancy or a type of B cell malignancy in the sample.

10

Detection of hybridization of RNA encoding IRTA proteins will indicate that a malignancy is a B cell malignancy. More specifically, detection of hybridization of RNA encoding ITRA1 protein indicates that the B cell 15 malignancy is a Mucosa-Associated-Lymphoid Tissue (MALT) B cell lymphoma. Detection of hybridization of RNA encoding ITRA4 and IRTA5 proteins indicate that the B cell malignancy is a mantle cell lymphoma. In an embodiment of the above-described method, the B cell malignancy 20 comprises a 1q21 chromosomal rearrangement. One of skill will use the above-described method as a diagnostic aid in conjunction with other standard methods of detecting/diagnosing malignancies, e.g. pathology of a tumor sample, which may indicate lymphoma and the above- 25 described method will then narrow the malignancy to a B cell lymphoma or more specifically to MALT) B cell lymphoma or a mantle cell lymphoma as discussed supra.

One of skill is familiar with known methods of detecting 30 of hybridization nucleic acid molecules to nucleic acid oligonucleotides, i.e. nucleic acid probes encoding a protein of interest for diagnostic methods. The nucleic

-37-

acid molecules encoding the IRTA proteins of the subject invention are useful for detecting B cell malignancy. One of skill will recognize that variations of the above-described method for detecting a B cell malignancy in a sample include, but are not limited to, digesting nucleic acid from the sample with restriction enzymes and separating the nucleic acid molecule fragments so obtained by size fractionation before hybridization.

In an embodiment of the above-described method for detecting a B cell malignancy in a sample from a subject, wherein the detectable marker is radioactive isotope, enzyme, dye, biotin, a fluorescent label or a chemiluminescent label. In a preferred embodiment, the B cell malignancy is selected from the group consisting of B cell lymphoma, multiple myeloma, Burkitt's lymphoma, marginal zone lymphoma, diffuse large cell lymphoma and follicular lymphoma cells. In a further embodiment, the B cell lymphoma is Mucosa-Associated-Lymphoid Tissue B cell lymphoma (MALT). In another preferred embodiment, the B cell lymphoma is non-Hodgkin's lymphoma.

This invention provides an antisense oligonucleotide having a sequence capable of specifically hybridizing to an mRNA molecule encoding a human ITRA protein so as to prevent overexpression of the mRNA molecule.

In preferred embodiments of the antisense oligonucleotide, the ITRA protein selected from the group consisting of human IRTA1, IRTA2, IRTA3, IRTA4 and IRTA5 protein. In further embodiments of any of the above-described oligonucleotides of nucleic acid molecules encoding the

-38-

IRTA1, IRTA2, IRTA3, IRTA4 and/or IRTA5 proteins, the nucleic acid may be genomic DNA or cDNA.

- One of skill is familiar with conventional techniques for  
5 nucleic acid hybridization of oligonucleotides, e.g. Ausubel, F.M. et al. *Current Protocols in Molecular Biology*, (John Wiley & Sons, New York, 1998), for example stringent conditions of 65°C in the presence of an elevated salt concentration. Such conditions are used for  
10 completely complementary nucleic acid hybridization, whereas conditions that are not stringent are used for hybridization of nucleic acids which are not totally complementary.
- 15 As used herein, the phrase "specifically hybridizing" means the ability of a nucleic acid molecule to recognize a nucleic acid sequence complementary to its own and to form double-helical segments through hydrogen bonding between complementary base pairs. As used herein, a  
20 "unique sequence" is a sequence specific to only the nucleic acid molecules encoding the IRTA1, IRTA2, IRTA3, IRTA4 and IRTA5 proteins. Nucleic acid probe technology is well known to those skilled in the art who will readily appreciate that such probes may vary greatly in length and  
25 may be labeled with a detectable label, such as a radioisotope or fluorescent dye, to facilitate detection of the probe. Detection of nucleic acid molecules encoding the IRTA1, IRTA2, IRTA3, IRTA4 and/or IRTA5 proteins is useful as a diagnostic test for any disease process in which levels of expression of the corresponding  
30 IRTA1, IRTA2, IRTA3, IRTA4 and/or IRTA5 proteins is altered. DNA probe molecules are produced by insertion of

-39-

- a DNA molecule which encodes mammalian IRTA1, IRTA2, IRTA3, IRTA4 and/or IRTA5 proteins or fragments thereof into suitable vectors, such as plasmids or bacteriophages, followed by insertion into suitable bacterial host cells
- 5 and replication and harvesting of the DNA probes, all using methods well known in the art. For example, the DNA may be extracted from a cell lysate using phenol and ethanol, digested with restriction enzymes corresponding to the insertion sites of the DNA into the vector
- 10 (discussed herein), electrophoresed, and cut out of the resulting gel. The oligonucleotide probes are useful for 'in situ' hybridization or in order to locate tissues which express this IRTA gene family, and for other hybridization assays for the presence of these genes
- 15 (nucleic acid molecules encoding any of the IRTA1-IRTA5 proteins) or their mRNA in various biological tissues. In addition, synthesized oligonucleotides (produced by a DNA synthesizer) complementary to the sequence of a DNA molecule which encodes an IRTA1, IRTA2, IRTA3, IRTA4 or
- 20 IRTA5 protein are useful as probes for these genes, for their associated mRNA, or for the isolation of related genes by homology screening of genomic or cDNA libraries, or by the use of amplification techniques such as the Polymerase Chain Reaction.
- 25
- This invention provides a purified IRTA1 protein comprising the amino acid sequence set forth in Figure 18A (SEQ ID NO:1). In an embodiment of the purified IRTA1 protein, wherein the IRTA1 protein is human IRTA1.
- 30
- This invention provides a purified IRTA2 protein comprising the amino acid sequence set forth in Figures

-40-

18B-1-18B-3 (SEQ ID NO:3). In an embodiment of the purified IRTA2 protein, the IRTA2 protein is human IRTA2.

This invention provides a purified IRTA3 protein  
5 comprising the amino acid sequence set forth in Figures  
18C-1-18C-2 (SEQ ID NO:5). In an embodiment of the purified IRTA3 protein, the IRTA3 protein is human IRTA3.

This invention provides a purified IRTA4 protein  
10 comprising the amino acid sequence set forth in Figures  
18D-1-18D-2 (SEQ ID NO: 7). In an embodiment of the purified IRTA3 protein, wherein the IRTA4 protein is human IRTA4.

15 This invention provides a purified IRTA5 protein comprising the amino acid sequence set forth in Figures  
18E-1-18E-2 (SEQ ID NO: 9). In an embodiment of the purified IRTA5 protein, the IRTA5 protein is human IRTA5.

20 In order to facilitate an understanding of the Experimental Details section which follows, certain frequently occurring methods and/or terms are best described in Sambrook, et al. (1989) and Harlow & Lane, Antibodies: A Laboratory Manual, Cold Spring Harbor  
25 Laboratories, Cold Spring Harbor, NY: 1988.

This invention provides an antibody/antibodies directed to an epitope of a purified IRTA1, IRTA2, IRTA3, IRTA4 or IRTA5 protein, or fragment(s) thereof, having the amino acid sequence set forth in any of Figures 18A, 18B-1-18B-3, 18C-1-18C-2, 18D-1-18D-2 or 18E-1-18E-2.

-41-

As used herein, the term "antibody" includes, but is not limited to, both naturally occurring and non-naturally occurring antibodies. Specifically, the term "antibody" includes polyclonal and monoclonal antibodies, and binding  
5 fragments thereof. Furthermore, the term "antibody" includes chimeric antibodies and wholly synthetic antibodies, and fragments thereof. The polyclonal and monoclonal antibodies may be "purified" which means the polyclonal and monoclonal antibodies are free of any other  
10 antibodies. As used herein, partially purified antibody means an antibody composition which comprises antibodies which specifically bind to any of the IRTA protein(s) of the subject invention, and consists of fewer protein impurities than does the serum from which the antibodies  
15 are derived. A protein impurity is a protein other than the antibodies specific for the IRTA protein(s) of the subject invention. For example, the partially purified antibodies may be an IgG preparation.  
  
20 Polyclonal antibodies (anti-IRTA antibodies) may be produced by injecting a host animal such as rabbit, rat, goat, mouse or other animal with the immunogen(s) of this invention, e.g. a purified human IRTA1, IRTA2, IRTA3, IRTA4 or IRTA5, described infra. The sera are extracted  
25 from the host animal and are screened to obtain polyclonal antibodies which are specific to the immunogen. Methods of screening for polyclonal antibodies are well known to those of ordinary skill in the art such as those disclosed in Harlow & Lane, Antibodies: A Laboratory Manual, (Cold  
30 Spring Harbor Laboratories, Cold Spring Harbor, NY: 1988)

-42-

the contents of which are hereby incorporated by reference.

The anti-IRTA monoclonal antibodies of the subject invention may be produced by immunizing for example, mice  
5 with an immunogen (the IRTA polypeptides or fragments thereof as described herein). The mice are inoculated intraperitoneally with an immunogenic amount of the above-described immunogen and then boosted with similar amounts of the immunogen. Spleens are collected from the  
10 immunized mice a few days after the final boost and a cell suspension is prepared from the spleens for use in the fusion.

Hybridomas may be prepared from the splenocytes and a murine tumor partner using the general somatic cell hybridization technique of Kohler, B. and Milstein, C., Nature (1975) 256: 495-497. Available murine myeloma lines, such as those from the American Type Culture Collection (ATCC), 10801 University Boulevard, Manassas,  
20 VA 20110-2209, USA, may be used in the hybridization. Basically, the technique involves fusing the tumor cells and splenocytes using a fusogen such as polyethylene glycol. After the fusion the cells are separated from the fusion medium and grown in a selective growth medium, such  
25 as HAT medium, to eliminate unhybridized parent cells. The hybridomas may be expanded, if desired, and supernatants may be assayed by conventional immunoassay procedures, for example radioimmunoassay, using the immunizing agent as antigen. Positive clones may be  
30 characterized further to determine whether they meet the criteria of the invention antibodies.

-43-

Hybridomas that produce such antibodies may be grown in vitro or in vivo using known procedures. The monoclonal antibodies may be isolated from the culture media or body fluids, as the case may be, by conventional immunoglobulin 5 purification procedures such as ammonium sulfate precipitation, gel electrophoresis, dialysis, chromatography, and ultrafiltration, if desired.

In the practice of the subject invention any of the above-10 described antibodies may be labeled with a detectable marker. In one embodiment, the labeled antibody is a purified labeled antibody. The term "antibody" includes, by way of example, both naturally occurring and non-naturally occurring antibodies. Specifically, the term 15 "antibody" includes polyclonal and monoclonal antibodies, and fragments thereof. Furthermore, the term "antibody" includes chimeric antibodies and wholly synthetic antibodies, and fragments thereof. A "detectable moiety" which functions as detectable labels are well known to 20 those of ordinary skill in the art and include, but are not limited to, a fluorescent label, a radioactive atom, a paramagnetic ion, biotin, a chemiluminescent label or a label which may be detected through a secondary enzymatic or binding step. The secondary enzymatic or binding step 25 may comprise the use of digoxigenin, alkaline phosphatase, horseradish peroxidase,  $\beta$ -galactosidase, fluorescein or streptavidin/biotin. Methods of labeling antibodies are well known in the art.

30 Methods of recovering serum from a subject are well known to those skilled in the art. Methods of partially purifying antibodies are also well known to those skilled

-44-

in the art, and include, by way of example, filtration, ion exchange chromatography, and precipitation.

The polyclonal and monoclonal antibodies of the invention  
5 may be labeled with a detectable marker. In one embodiment, the labeled antibody is a purified labeled antibody. The detectable marker may be, for example, a radioactive or fluorescent marker. Methods of labeling antibodies are well known in the art.

10

Determining whether the polyclonal and monoclonal antibodies of the subject invention bind to cells, e.g. cancer cells, expressing an IRTA protein and form a complex with one or more of the IRTA protein(s) described  
15 herein, or fragments thereof, on the surface of said cells, may be accomplished according to methods well known to those skilled in the art. In the preferred embodiment, the determining is accomplished according to flow cytometry methods.

20

The antibodies of the subject invention may be bound to an insoluble matrix such as that used in affinity chromatography. Cells which form a complex, i.e. bind, with the immobilized polyclonal or monoclonal antibody may  
25 be isolated by standard methods well known to those skilled in the art. For example, isolation may comprise affinity chromatography using immobilized antibody.

Alternatively, the antibody may be a free antibody. In  
30 this case, isolation may comprise cell sorting using free, labeled primary or secondary antibodies. Such cell sorting

-45-

methods are standard and are well known to those skilled in the art.

This invention provides an antibody directed to a purified  
5 IRTA protein selected from the group consisting of IRTA1,  
IRTA2, IRTA3, IRTA4 and IRTA5. In a preferred embodiment  
of the anti-IRTA antibody the IRTA protein is human IRTA  
protein. The IRTA protein may be any mammalian IRTA  
protein, including a murine IRTA protein. In a further  
10 embodiment of any the above-described antibodies, the  
antibody is a monoclonal antibody. In another embodiment,  
the monoclonal antibody is a murine monoclonal antibody or  
a humanized monoclonal antibody. As used herein,  
"humanized" means an antibody having characteristics of a  
15 human antibody, such antibody being non-naturally  
occurring, but created using hybridoma techniques wherein  
the antibody is of human origin except for the antigen  
determinant portion, which is murine. In yet another  
embodiment, the antibody is a polyclonal antibody.

20

In preferred embodiments, any of the antibodies of the  
subject invention may be conjugated to a therapeutic  
agent. In further preferred embodiments, the therapeutic  
agent is a radioisotope, toxin, toxoid, or  
25 chemotherapeutic agent. The conjugated antibodies of the  
subject invention may be administered to a subject having  
a B cell cancer in any of the methods provided below.

This invention provides a pharmaceutical composition  
30 comprising an amount of the antibody directed to an IRTA  
protein effective to bind to cancer cells expressing an  
IRTA protein selected from the group consisting of human

-46-

IRTA1, IRTA2, IRTA3, IRTA4 and IRTA5 so as to prevent growth of the cancer cells and a pharmaceutically acceptable carrier. The anti-IRTA antibody may be directed to an epitope of an IRTA protein selected from the group 5 consisting of IRTA1, IRTA2, IRTA3, IRTA4 and IRTA5. The IRTA proteins may be human or mouse IRTA proteins.

In preferred embodiments of the above-described pharmaceutical composition, the cancer cells are selected 10 from the group consisting of B cell lymphoma, multiple myeloma, a mantle cell lymphoma, Burkitt's lymphoma, marginal zone lymphoma, diffuse large cell lymphoma and follicular lymphoma cells. In another preferred embodiment of the pharmaceutical composition, the B cell 15 lymphoma cells are Mucosa-Associated-Lymphoid Tissue B cell lymphoma (MALT) cells. In another preferred embodiment of the pharmaceutical composition, the B cell lymphoma cells are non-Hodgkin's lymphoma cells.

20 This invention provides a pharmaceutical composition comprising an amount of any of the above-described oligonucleotides effective to prevent overexpression of a human IRTA protein and a pharmaceutically acceptable carrier capable. In preferred embodiments of the 25 pharmaceutical composition the oligonucleotide is a nucleic acid molecule which encodes an IRTA protein selected from the group consisting of IRTA1, IRTA2, IRTA3, IRTA4 and IRTA5. The IRTA proteins may be human or mouse IRTA proteins.

30

As used herein, "malignant" means capable of metastasizing. As used herein, "tumor cells" are cells

-47-

which originate from a tumor, i.e., from a new growth of different or abnormal tissue. The tumor cells and cancer cells may exist as part of the tumor mass, or may exist as free-floating cells detached from the tumor mass from 5 which they originate.

As used herein, malignant cells include, but are in no way limited to, B cell lymphoma, multiple myeloma, Burkitt's lymphoma, mantle cell lymphoma, marginal zone lymphoma, 10 diffuse large cell lymphoma and follicular lymphoma. The B cell lymphoma is Mucosa-Associated-Lymphoid Tissue B cell lymphoma (MALT) or is non-Hodgkin's lymphoma.

As used herein, "subject" is any animal or artificially modified animal. Artificially modified animals include, but are not limited to, SCID mice with human immune systems. In a preferred embodiment, the subject is a 15 human.

20 This invention provides a method of diagnosing B cell malignancy which comprises a 1q21 chromosomal rearrangement in a sample from a subject which comprises: a) obtaining the sample from the subject; b) contacting the sample of step (a) with an antibody directed to a 25 purified IRTA protein capable of specifically binding with a human IRTA protein selected from the group consisting of human IRTA1, IRTA2, IRTA3, IRTA4 and IRTA5 IRTA protein on a cell surface of a cancer cell under conditions permitting binding of the antibody with human IRTA protein 30 on the cell surface of the cancer cell, wherein the antibody is labeled with a detectable marker; and c) detecting any binding in step (b), wherein detection of

-48-

binding indicates a diagnosis of B cell malignancy in the sample.

In an embodiment of the above-described method of  
5 diagnosing B cell malignancy, the IRTA protein is selected  
from the group consisting of IRTA1, IRTA2, IRTA3, IRTA4  
and IRTA5. In another embodiment of the method the IRTA  
protein is human or mouse IRTA protein. In a further  
embodiment IRTA protein is purified. In a preferred  
10 embodiment of this method, the B cell malignancy is  
selected from the group consisting of B cell lymphoma,  
multiple myeloma, Burkitt's lymphoma, marginal zone  
lymphoma, diffuse large cell lymphoma and follicular  
lymphoma. In yet another embodiment of this method, the B  
15 cell lymphoma is Mucosa-Associated-Lymphoid Tissue B cell  
lymphoma (MALT). In another preferred embodiment of this  
method, the B cell lymphoma is non-Hodgkin's lymphoma.

This invention provides a method of detecting human IRTA  
20 protein in a sample which comprises: a) contacting the  
sample with any of any of the above-described anti-IRTA  
antibodies under conditions permitting the formation of a  
complex between the antibody and the IRTA in the sample;  
and b) detecting the complex formed in step (a), thereby  
25 detecting the presence of human IRTA in the sample. In an  
embodiement the IRTA protein detected may be an IRTA1,  
IRTA2, IRTA3, IRTA4 or IRTA5 protein, having an amino acid  
sequence set forth in any of Figures 18A, 18B-1-18B-3,  
18C-1-18C-2, 18D-1-18D-2 or 18E-e-18E-2. As described  
30 hereinabove detection of the complex formed may be  
achieved by using antibody labeled with a detectable  
marker and determining presence of labeled complex.

-49-

Detecting human IRTA protein in a sample from a subject is another method of diagnosing B cell malignancy in a subject. In an embodiment of this method of diagnosis, the B cell malignancy is selected from the group consisting of  
5 B cell lymphoma, multiple myeloma, Burkitt's lymphoma, marginal zone lymphoma, diffuse large cell lymphoma and follicular lymphoma. In yet another embodiment of this method, the B cell lymphoma is Mucosa-Associated-Lymphoid Tissue B cell lymphoma (MALT). In another preferred  
10 embodiment of this method, the B cell lymphoma is non-Hodgkin's lymphoma.

This invention provides a method of treating a subject having a B cell cancer which comprises administering to  
15 the subject an amount of anti-IRTA antibody effective to bind to cancer cells expressing an IRTA protein so as to prevent growth of the cancer cells and a pharmaceutically acceptable carrier, thereby treating the subject. Growth and proliferation of the cancer cells is thereby inhibited  
20 and the cancer cells die. In an embodiment of the above-described method, the IRTA protein is selected from the group consisting of human IRTA1, IRTA2, IRTA3, IRTA4 and IRTA5. In a preferred embodiment of the above-described method of treating a subject having a B cell cancer, the  
25 anti-IRTA antibody is a monoclonal antibody. In another embodiment of the method, the monoclonal antibody is a murine monoclonal antibody or a humanized monoclonal antibody. The antibody may be a chimeric antibody. In a further embodiment, the anti-IRTA antibody is a polyclonal antibody. In an embodiment, the polyclonal antibody may be a murine or human polyclonal antibody. In  
30 a preferred embodiment, the B cell cancer is selected from

-50-

the group consisting of B cell lymphoma, multiple myeloma, Burkitt's lymphoma, mantle cell lymphoma marginal zone lymphoma, diffuse large cell lymphoma and follicular lymphoma. In another preferred embodiment, the B cell 5 lymphoma is Mucosa-Associated-Lymphoid Tissue B cell lymphoma (MALT). In a further preferred embodiment, the B cell lymphoma is non-Hodgkin's lymphoma. In a preferred embodiment of the above-described method of treating a subject having a B cell cancer, administration of the 10 amount of anti-IRTA antibody effective to bind to cancer cells expressing an IRTA protein is intravenous, intraperitoneal, intrathecal, intralymphatic, intramuscular, intralesional, parenteral, epidural, subcutaneous; by infusion, liposome-mediated delivery, 15 aerosol delivery; topical, oral, nasal, anal, ocular or otic delivery. In another preferred embodiment of the above-described methods, the anti-IRTA antibody may be conjugated to a therapeutic agent. In further preferred embodiments, the therapeutic agent is a radioisotope, 20 toxin, toxoid, or chemotherapeutic agent.

This invention provides a method of treating a subject having a B cell cancer which comprises administering to the subject an amount of an antisense oligonucleotide 25 having a sequence capable of specifically hybridizing to an mRNA molecule encoding a human ITRA protein so as to prevent overexpression of the human ITRA protein, so as to arrest cell growth or induce cell death of cancer cells expressing ITRA protein(s) and a pharmaceutically acceptable carrier, thereby treating the subject.

-51-

In an embodiment of the above-described method of treating a subject having a B cell cancer, the IRTA protein is selected from the group consisting of human IRTA1, IRTA2, IRTA3, IRTA4 and IRTA5 protein. In a preferred embodiment, B cell cancer is selected from the group consisting of B cell lymphoma, multiple myeloma, Burkitt's lymphoma, marginal zone lymphoma, diffuse large cell lymphoma and follicular lymphoma. In another preferred embodiment, the B cell lymphoma is Mucosa-Associated-Lymphoid Tissue-B cell lymphoma (MALT). In a yet another preferred embodiment, the B cell lymphoma is non-Hodgkin's lymphoma. In embodiments of any of the above-described oligonucleotides of nucleic acid molecules encoding the IRTA1, IRTA2, IRTA3, IRTA4 and/or IRTA5 proteins, the nucleic acid may be genomic DNA or cDNA. In a further preferred embodiment of the above-described method of treating a subject having a B cell cancer, administration of the amount of oligonucleotide of effective to prevent overexpression of a human IRTA protein is intravenous, intraperitoneal, intrathecal, intralymphatic, intramuscular, intralesional, parenteral, epidural, subcutaneous; by infusion, liposome-mediated delivery, aerosol delivery; topical, oral, nasal, anal, ocular or otic delivery. In an embodiment of the above-described methods, the oligonucleotide may be conjugated to a therapeutic agent. In further preferred embodiments, the therapeutic agent is a radioisotope, toxin, toxoid, or chemotherapeutic agent.

The invention also provides a pharmaceutical composition comprising either an effective amount of the

-52-

oligonucleotides or of the antibodies described above and a pharmaceutically acceptable carrier. In the subject invention an "effective amount" is any amount of an oligonucleotide or an antibody which, when administered to 5 a subject suffering from a disease or abnormality against which the oligonucleotide or antibody are effective, causes reduction, remission, or regression of the disease or abnormality. In the practice of this invention the "pharmaceutically acceptable carrier" is any physiological 10 carrier known to those of ordinary skill in the art useful in formulating pharmaceutical compositions.

Pharmaceutically acceptable carriers are well known to those skilled in the art and include, but are not limited 15 to, 0.01-0.1M and preferably 0.05M phosphate buffer or 0.8% saline. Additionally, such pharmaceutically acceptable carriers may be aqueous or non-aqueous solutions, suspensions, and emulsions. Examples of non-aqueous solvents are propylene glycol, polyethylene 20 glycol, vegetable oils such as olive oil, and injectable organic esters such as ethyl oleate. Aqueous carriers include water, alcoholic/aqueous solutions, emulsions or suspensions, including saline and buffered media. Parenteral vehicles include sodium chloride solution, 25 Ringer's dextrose, dextrose and sodium chloride, lactated Ringer's or fixed oils. Intravenous vehicles include fluid and nutrient replenishers, electrolyte replenishers such as those based on Ringer's dextrose, and the like. Preservatives and other additives may also be present, 30 such as, for example, antimicrobials, antioxidants, chelating agents, inert gases and the like.

-53-

In one preferred embodiment the pharmaceutical carrier may be a liquid and the pharmaceutical composition would be in the form of a solution. In another equally preferred embodiment, the pharmaceutically acceptable carrier is a 5 solid and the composition is in the form of a powder or tablet. In a further embodiment, the pharmaceutical carrier is a gel and the composition is in the form of a suppository or cream. In a further embodiment the compound may be formulated as a part of a pharmaceutically 10 acceptable transdermal patch.

A solid carrier can include one or more substances which may also act as flavoring agents, lubricants, solubilizers, suspending agents, fillers, glidants, 15 compression aids, binders or tablet-disintegrating agents; it can also be an encapsulating material. In powders, the carrier is a finely divided solid which is in admixture with the finely divided active ingredient. In tablets, the active ingredient is mixed with a carrier having the 20 necessary compression properties in suitable proportions and compacted in the shape and size desired. The powders and tablets preferably contain up to 99% of the active ingredient. Suitable solid carriers include, for example, calcium phosphate, magnesium stearate, talc, sugars, 25 lactose, dextrin, starch, gelatin, cellulose, polyvinylpyrrolidine, low melting waxes and ion exchange resins.

Liquid carriers are used in preparing solutions, 30 suspensions, emulsions, syrups, elixirs and pressurized compositions. The active ingredient can be dissolved or suspended in a pharmaceutically acceptable liquid carrier

-54-

such as water, an organic solvent, a mixture of both or pharmaceutically acceptable oils or fats. The liquid carrier can contain other suitable pharmaceutical additives such as solubilizers, emulsifiers, buffers, 5 preservatives, sweeteners, flavoring agents, suspending agents, thickening agents, colors, viscosity regulators, stabilizers or osmo-regulators. Suitable examples of liquid carriers for oral and parenteral administration include water (partially containing additives as above, 10 e.g. cellulose derivatives, preferably sodium carboxymethyl cellulose solution), alcohols (including monohydric alcohols and polyhydric alcohols, e.g. glycols) and their derivatives, and oils (e.g. fractionated coconut oil and arachis oil). For parenteral administration, the 15 carrier can also be an oily ester such as ethyl oleate and isopropyl myristate. Sterile liquid carriers are useful in sterile liquid form compositions for parenteral administration. The liquid carrier for pressurized compositions can be halogenated hydrocarbon or other 20 pharmaceutically acceptable propellant.

Liquid pharmaceutical compositions which are sterile solutions or suspensions can be utilized by for example, intramuscular, intrathecal, epidural, intraperitoneal or 25 subcutaneous injection. Sterile solutions can also be administered intravenously. The compounds may be prepared as a sterile solid composition which may be dissolved or suspended at the time of administration using sterile water, saline, or other appropriate sterile injectable 30 medium. Carriers are intended to include necessary and inert binders, suspending agents, lubricants, flavorants, sweeteners, preservatives, dyes, and coatings.

-55-

- The pharmaceutical composition comprising the oligonucleotide or the antibody can be administered orally in the form of a sterile solution or suspension containing other solutes or suspending agents, for example, enough  
5 saline or glucose to make the solution isotonic, bile salts, acacia, gelatin, sorbitan monoleate, polysorbate 80 (oleate esters of sorbitol and its anhydrides copolymerized with ethylene oxide) and the like.
- 10 The pharmaceutical composition comprising the oligonucleotide or the antibody can also be administered orally either in liquid or solid composition form. Compositions suitable for oral administration include solid forms, such as pills, capsules, granules, tablets,  
15 and powders, and liquid forms, such as solutions, syrups, elixirs, and suspensions. Forms useful for parenteral administration include sterile solutions, emulsions, and suspensions.
- 20 Optimal dosages to be administered may be determined by those skilled in the art, and will vary with the particular inhibitor in use, the strength of the preparation, the mode of administration, and the advancement of the disease condition or abnormality.  
25 Additional factors depending on the particular subject being treated will result in a need to adjust dosages, including subject age, weight, gender, diet, and time of administration.
- 30 This invention will be better understood from the Experimental Details which follow. However, one skilled in the art will readily appreciate that the specific

-56-

methods and results discussed are merely illustrative of the invention as described more fully in the claims which follow thereafter.

-57-

#### EXPERIMENTAL DETAILS

##### First Series of Experiments

Molecular analysis of chromosomal translocations associated with multiple myeloma (MM) has indicated that the pathogenesis of this malignancy may be heterogeneous, being associated with several distinct oncogenes including BCL-1, MUM-1 and FGFR3. Structural abnormalities of chromosome 1q21, including translocations with chromosome 14q32, represent frequent cytogenetic aberrations associated with multiple myeloma. In order to identify the genes involved in these translocations, the breakpoint regions corresponding to both derivatives of a t(1;14)(q21;q32) detectable in the FR4 human plasmacytoma cell line were cloned. Analysis of the breakpoint sequences showed that they involved a reciprocal recombination between the Immunoglobulin heavy chain (IgH) locus on 14q32 and unknown sequences on 1q21. The normal locus corresponding to the 1q21 region involved in the translocation was cloned and the genes adjacent to the breakpoint region were identified by an exon-trapping strategy. Two genes were found, located within a 20 Kb distance from each other, in the region spanning the breakpoint on 1q21. The first gene, called MUM-2 (multiple myeloma-2) is expressed as a 2.5 Kb mRNA transcript detectable in spleen and lymph nodes. Cloning and sequencing of the full-length MUM-2 cDNA predicts a 515 amino acid cell surface glycoprotein containing four extracellular Ig-type domains, a transmembrane and a cytoplasmic domain and sharing a 37% identity (51% homology) with Fc gamma receptor I over its first three extracellular domains. In FR4 cells, the

-58-

translocation breakpoints interrupt the MUM-2 coding domain and juxtapose it to the IgH locus in the same transcriptional orientation. As a consequence, structurally abnormal FR4-specific MUM-2 transcripts (3.0, 5 5.2 and 6.0 Kb) in lymph nodes and spleen and encodes a protein with an extracellular domain containing six Ig-type domains homologous to members of the Fc gamma and Ig-type adhesion receptor families. The structure of the MUM-2 and MUM-3 genes and their direct involvement in a 10 MM-associated translocation suggest that these genes code for novel cell surface receptors important for normal lymphocyte function and B cell malignancy.

Second Series of Experiments

15 **Experimental Procedures**

**Cell Lines**

The MM cell lines used in this study (FR4, U266, JJN3, EJM, SKMM1, RPMI-8226, XG1, XG2, XG4, XG6, XG7) have been previously reported (Tagawa et al., 1990), (Jernberg et 20 al., 1987), (Hamilton et al., 1990; Jackson et al., 1989), (Eton et al., 1989), (Zhang et al., 1994). The FR4 cell line was established in the laboratory of one of the authors (S.T.). The U266, JJN3, and EJM cell lines were gifts from Dr. K. Nilsson (University of Uppsala, Uppsala, 25 Sweden) and the SKMM-1 cell line was a gift of A.N. Houghton (Memorial Sloan Kettering Cancer Center, New York, NY). The five XG cell lines were obtained from Dr. Bernard Klein and cultured in the presence of 1 ng/ml human recombinant IL-6 as described previously (Zhang et 30 al., 1994). The BL cell lines with 1q21 abnormalities

-59-

have been previously described (Polito et al., 1995); (Magrath et al., 1980) and were grown in RPMI, 10% FCS.

5           **Genomic and cDNA library screening and DNA sequence analysis**

Two genomic libraries were constructed from FR4 genomic DNA either by BamHI complete digestion or by Sau3AI partial digestion and subsequent ligation of gel-purified 10 fractions into the 1DASH-II phage vector (Stratagene). The BamHI library was screened with a 4.2 kb XhoI-BamHI probe derived from the Ca locus and the Sau3AI library was screened with a 5'Sa probe previously described (Bergsagel et al., 1996). A human placental DNA library (Stratagene) 15 was screened with probe 1.0EH (Figures 8A-8C) to obtain the germline 1q21 locus. Library screening and plaque isolation were preformed according to established procedures (Sambrook et al., 1989). *IRTA1* and *IRTA2* cDNA clones were isolated from an oligo-dT/random-primed cDNA 20 library constructed from normal human spleen RNA (Clontech). The *IRTA1* cDNA probe used for library screening was obtained from RT-PCR of human spleen cDNA using primers flanking exons 1 and 3. DNA sequencing was preformed on an ABI 373 automated sequencer (Applied 25 Biosystems). Sequence homology searches were carried out through the BLAST e-mail server at the National Center for Biotechnology Information, Bethesda, MD.

30           **PAC and YAC isolation and exon trapping**

Human PAC clones were obtained by screening a human PAC library spotted onto nylon membranes (Research Genetics), with the 1.0 EH probe (Figures 8A-8C). The Zeneca

-60-

(formerly ICI) human YAC library (Anand et al., 1990) obtained from the United Kingdom Human Genome Mapping Resource Center (UK-HGMP) was screened using a PCR-based pooling strategy. Exon trapping was performed using the 5 exon trapping system (Gibco BRL), according to the manufacturer's instructions.

10 Isolation of PAC/YAC end clones, pulsed-field gel electrophoresis (PFGE) and fluorescence *in situ* hybridization (FISH) analysis

PAC DNA extraction was performed according to standard alkaline lysis methods (Drakopoli N et al., 1996). A vectorette-PCR method was used to isolate PAC and YAC end 15 probes (Riley et al., 1990), as previously described (Iida et al, 1996). PFGE analysis was performed according to standard protocols (Drakopoli N et al., 1996) using the CHEF Mapper system (BioRad, Hercules, CA). Biotin labeling of PAC DNA, chromosome preparation and FISH were 20 performed as previously described (Rao et al., 1993).

25 Southern and Northern blot analyses, RACE and RT-PCR  
Southern and northern blot analyses were performed as described previously (Neri et al, 1991). For Northern blot analyses total RNA was prepared by the guanidium thiocyanate method and poly(A) RNA was selected using poly(T)-coated beads (Oligotex Kit by Qiagen). For Northern blots, 2 mg of poly(A) RNA were loaded per lane. Multiple tissue Northern filters were obtained from 30 Clontech. RACE was performed using the Marathon cDNA Amplification kit (Clontech) and Marathon-Ready spleen

-61-

cDNA. First strand cDNA synthesis was performed using the Superscript RT-PCR system (Gibco BRL)

*In situ hybridization*

5 Digoxigenin-containing antisense and sense cRNA probes were transcribed with T3 and T7 RNA polymerase, respectively, from linearized pBluescript KS+ plasmids containing coding region of cDNAs (nucleotides 62 to 1681 of IRTA1 and 18 to 2996 of IRTA2.) Hyperplastic human  
10 tonsillar tissue surgically resected from children in Babies' Hospital, Columbia Presbyterian Medical Center was snap frozen in powdered dry ice. Cryostat sections were stored for several days at -80 degrees C prior to processing. Non-radioactive *in situ* hybridization was  
15 performed essentially as described (Frank et al., 1999), except that fixation time in 4% paraformaldehyde was increased to 20 minutes, and proteinase K treatment was omitted. The stringency of hybridization was 68 degrees C, in 5X SSC, 50% formamide. Alkaline phosphatase-  
20 conjugated anti-digoxigenin antibody staining was developed with BCIP/NBT substrate.

*Transfection, immunoprecipitation and Western Blotting*

25 293 cells (ATCC), grown in DMEM, 10% FCS were transiently transfected, according to the standard calcium phosphate method, with pMT2T and pMT2T-IRTA1/Ca transient expression constructs. The latter was generated using the IRTA1/Ca RT-PCR product from FR4. Cells ( $2 \times 10^6$  of transfectants and  $2 \times 10^7$  of remaining cell lines) were  
30 solubilized in Triton X-100 lysis buffer (150 mM NaCl, 10 mM Tris-HCl [pH 7.4], 1% Tx-100, 0.1% BSA) in the presence of a protease inhibitors coctail (Roche Biochemicals).

-62-

Lysates were incubated at 4°C for 2 hours with 4 mg/ml of the monoclonal antibody #117-332-1 (Yu et al., 1990) (Tanox Biosystems, Inc, Houston, Texas) that was raised against the extracellular portion of the IgA membrane peptide. Immune complexes were isolated with protein G-Sepharose (Pharmacia) prior to electrophoresis on 10-20% Tris-HCl gradient gels (Biorad) and immunoblotting, using 15 mg/ml of the #117-332-1 antibody. Results were visualized by ECL (Amersham).

10

## RESULTS

### Molecular Cloning of the t(1;14) (q21;q32)

Chromosomal translocations involving the Ig heavy-chain (IGH) locus often occur within or near IgH switch regions as a result of "illegitimate" switch recombination events (Dalla-Favera et al., 1983; Chesi et al., 1996; Chesi et al., 1998). The breakpoints can be detected by Southern-blot hybridization assays as rearranged alleles in which the IGH constant ( $C_H$ ) region sequences have lost their syntenic association with IGH joining ( $J_H$ ) and 5' switch region (S) sequences (Dalla-Favera et al., 1983; Neri et al., 1988; Neri et al., 1991; Bergsagel et al., 1996). This assay has led to the identification of several chromosomal partners for the IgH locus in B-NHL and MM (Taub et al., 1982; Dalla-Favera et al., 1983; Neri et al., 1988; Neri et al., 1991; Ye et al., 1993; Chesi et al., 1996; Richelda et al., 1997; Iida et al., 1997; Dyomin et al., 1997; Dyomin et al., 2000). We employed the same strategy in order to clone the 1q21 breakpoint region in FR4, a myeloma cell line carrying a t(1;14) (q21;q32), as determined by cytogenetic analysis.

-63-

(Tagawa et al., 1990; Taniwaki M, unpublished results). Two "illegitimately" rearranged fragments were identified within the  $\text{C}\alpha$  heavy-chain locus in FR4 by Southern blot hybridization analysis (data not shown), and were cloned from phage libraries constructed from FR4 genomic DNA. Restriction mapping, Southern blot hybridization and partial nucleotide sequencing of two genomic phages (clones  $\lambda$  FR4B-5 and  $\lambda$  FR4S-a, Figure 8A) demonstrated that they contained the chromosomal breakpoints of a reciprocal balanced translocation between the  $\text{C}\alpha_1$  locus on 14q32 and non-IGH sequences. A probe (1.0EH) representing these non-IgH sequences (Figure 8A) was then used to clone the corresponding normal genomic locus from phage, P1 artificial chromosome (PAC), and yeast artificial chromosome (YAC) human genomic libraries. Fluorescence *in situ* hybridization (FISH) analysis of normal human metaphase spreads using the 100-kb non-chimaeric PAC clone 49A16 which spans the breakpoint region (see below, Figure 13), identified the partner chromosomal locus as derived from band 1q21 (Figure 8C). Mapping to a single locus within chromosome 1 was confirmed by hybridization of two non-repetitive probes to DNA from a somatic-cell hybrid panel representative of individual human chromosomes (data not shown). These results were consistent with the cloning of sequences spanning the t(1;14)(q21;q32) in FR4.

Sequence analysis of the breakpoint regions on the derivative chromosomes and alignment with the germline 14q32 and 1q21 loci revealed that the breakpoint had occurred in the intron between the CH3 and the

-64-

transmembrane exon of  $\text{Ca}_1$  on chromosome 14. Although the breakpoint region was devoid of recombination signal sequences (RSS) or switch signal sequences (Kuppers et al., 1999), the sequence CTTAAC (underlined on Figure 8B) was present in both germline chromosomes 14 and 1 at the breakpoint junction. One copy of this sequence was present in each of the derivative chromosomes, with a slight modification in the der(1) copy (point mutation in the last nucleotide: C to G). The nucleotides AT preceding CTTAAC on chromosome 1 were also present in both derivative chromosomes (Figure 8B). The translocation did not result in any loss of chromosome 1 sequences. On the other hand, in the chromosome 14 portion of der(1) we observed two deletions upstream of the breakpoint junction: a 16 nucleotide deletion (GGCACCTCCCCTTAAC) and a 4 nucleotide deletion (TGCA). 6 nucleotides upstream (Figure 8B). These observations indicate that the t(1;14)(q21;q32) in FR4 cells represents a balanced reciprocal translocation possibly facilitated by the presence of homologous sequences.. (CTTAAC) on both chromosomes.

The 1q21 breakpoint region contains genes coding for novel members of the Immunoglobulin Receptor Superfamily

We next investigated whether the region of chromosome 1q21 spanning the translocation breakpoint in FR4 contains a transcriptional unit. DNA from partially overlapping PAC clones 49A16 and 210K22 (Figure 13) was "shotgun" cloned in plasmids, sequenced and analyzed for homology to known genes in human genome databases. In parallel, candidate

-65-

genes on the 49A16 PAC were sought by an exon trapping strategy (Church et al., 1994).

Mapping of the candidate exons on the 1q21 genomic clones 5 revealed that the FR4 breakpoint had occurred between two trapped exons (see below, Figure 13), which belonged to the same transcript since they could be linked by RT-PCR using spleen RNA. This RT-PCR product was then used as a probe to screen a spleen cDNA library in order to isolate 10 full-length clones corresponding to this transcript. Two sets of cDNA clones were identified, belonging to two distinct transcripts and sharing a 76% mRNA sequence identity within the 443 bp probe region. Full length cDNA clones for both transcripts were obtained by rapid 15 amplification of cDNA ends (RACE) on human spleen cDNA that generated 5' and 3' extension products.

The schematic structure of the cDNA representing the first transcript is depicted in Figure 9A. Alternate usage of 20 three potential polyadenylation sites in its 3' untranslated region gives rise to three mRNA species of 2.6, 2.7 and 3.5 kb, encoding the same putative 515-amino acid protein (Figure 9A). The predicted features of this protein include a signal peptide, in accordance with the 25 [-3, -1] rule (von Heijne, 1986), four extracellular Ig-type domains carrying three potential asparagine (N)-linked glycosylation sites (Figure 9A), a 16 amino acid transmembrane and a 106 amino acid cytoplasmic domain with three putative consensus Src-homology 2 (SH2)-binding 30 domains (Unkeless and Jin, 1997) (Figure 10B). These (SH2)-binding domains exhibit features of both ITAM (Immune-receptor Tyrosine-based Activation Motif

-66-

-D/EX,D/EX<sub>2</sub>YXXL/IX<sub>6-8</sub>YXXL/I; where X denotes non-conserved residues) (Reth, 1989) and ITIM motifs (Immune-receptor Tyrosine-based Inhibition Motif - S/V/L/IYXXL/V where X denotes non-conserved residues) (Unkeless and Jin, 1997).

5 As shown in Figure 10B, the first two SH2-binding domains are spaced 8 aminoacids apart, consistent with the consensus ITAM motif. Diverging from the consensus, the glutamate residue (E) is positioned four rather than two aminoacids before the first tyrosine (Y) (Figure 10B), and

10 the +3 position relative to tyrosine (Y) is occupied by valine (V) rather than leucine (L) or isoleucine (I) (Cambier, 1995). All three domains conform to the ITIM consensus and each is encoded by a separate exon, as is the case for ITIM. Thus their arrangement may give rise

15 to three ITIM or possibly to one ITAM and one ITIM. The overall structure of this protein suggests that it represents a novel transmembrane receptor of the Ig superfamily and it was therefore named IRTA1 (Immune Receptor Translocation Associated gene 1).

20

The second cDNA shares homology to IRTA1 (68% nucleotide identity for the length of the IRTA1 message encoding its extracellular domain) and was named IRTA2. The IRTA2 locus is more complex than IRTA1 and is transcribed into

25 three major mRNA isoforms (IRTA2a, IRTA2b, IRTA2c) of different molecular weight (2.8, 4.7 and 5.4 kb respectively), each with its own unique 3' untranslated region (Figure 9B). In addition, a 0.6 kb transcript (Figure 12A) arises from the usage of an early

30 polyadenylation signal at nucleotide 536 of IRTA2. The three predicted IRTA2 protein isoforms encoded by these

-67-

transcripts share a common aminoacid sequence until residue 560, featuring a common signal peptide and six extracellular Ig-type domains (Figure 9B). IRTA2a encodes for a 759 aa secreted glycoprotein with eight Ig-type domains followed by 13 unique, predominantly polar aminoacids at its C-terminus. IRTA2b diverges from IRTA2a at amino acid residue 560, and extends for a short stretch of 32 additional residues, whose hydrophobicity is compatible with its docking to the plasma membrane via a GPI-anchor (Ferguson and Williams, 1988). IRTA2c is the longest isoform whose sequence deviates from IRTA2a at aminoacid 746. It encodes a 977 aa type I transmembrane glycoprotein with nine extracellular Ig-type domains, harboring eight potential N-linked glycosylation sites, a 23 aminoacid transmembrane and a 104 aminoacid cytoplasmic domain with three consensus SH2-binding motifs (Figure 10B). Each of the SH2-binding sites in IRTA2c agrees with the ITIM consensus (Figure 10B) and is encoded by a separate exon. These features suggest that IRTA2c is a novel transmembrane receptor of the Ig superfamily with secreted and GPI-linked isoforms.

**Homology between the IRTA proteins and Immunoglobulin Superfamily Receptors**

25

Amino acid alignment of the entire extracellular domains of the IRTA1 and IRTA2 proteins to each other and to other Ig superfamily members revealed a remarkable homology between them (47% identity and 51% similarity) and a lower, but striking homology to the Fc gamma receptor family of proteins. This homology was stronger in the aminoacid positions conserved among the different classes

-68-

of Fc receptors. Among Fc receptors, the high affinity IgG receptor FCGRI (CD64) shared the highest levels of homology with the first three Ig-domains of IRTA1 and IRTA2 (37% identity and 50% similarity) throughout its entire extracellular portion (Figure 10A). Lower levels of homology were observed between the IRTA proteins and the extracellular domains of other cell surface molecules, including human platelet endothelial cell adhesion molecule (PECAM1), B-lymphocyte cell adhesion molecule (CD22) and Biliary Glycoprotein 1 (BGP1) (22-25% identity, 38-41% homology).

No homology is apparent between the IRTAs and members of the Fc receptor family in their cytoplasmic domains. In contrast, significant aminoacid homology is present between IRTA1 and PECAM1 (31% aminoacid identity and 45% homology), IRTA2c and BGP1 (30% identity, 35% homology) and IRTA2c and PECAM1 (28% identity, 50% homology) (Figure 10B). These homologies suggest employment of similar downstream signaling pathways by these different proteins.

IRTA1 and IRTA2 are normally expressed in specific subpopulations of B cells

The normal expression pattern of the IRTA1 and IRTA2 mRNAs was first analyzed by Northern blot hybridization of RNA derived from different normal human tissues and from human cell lines representing different hematopoietic lineages and stages of B-cell development.

-69-

*IRTA1* expression was detected at a very low level in human spleen and lymph node RNA (Figure 11A, left panel) and was undetectable in all other human tissues analyzed, including fetal liver, bone marrow, lung, placenta, small intestine, kidney, liver, colon, skeletal muscle, heart and brain (data not shown). Among B cell lines, *IRTA1* expression was absent in cell lines representing pre-B and germinal center B-cells, plasma cells and cells of erythroid, T-cell and myeloid origin (data not shown, see 10 Materials and Methods). Expression was detectable at very low levels only in EBV-immortalized lymphoblastoid cell lines (LCL), which represent a subpopulation (immunoblasts) positioned downstream of germinal center B cells in B-cell differentiation. However, expression was induced in estrogen-deprived ER/EB cells which, being immortalized by a recombinant EBV genome in which the EBNA2 gene is fused to the estrogen receptor, proliferate in the presence of estrogen while they arrest in the G<sub>0</sub>/G<sub>1</sub> phase upon estrogen deprivation (Kempkes et al., 1995). 15 20 *IRTA1* expression was barely detectable in these cells in the presence of estrogen, but was induced (10-fold) upon their G<sub>0</sub>/G<sub>1</sub> arrest following estrogen withdrawal (Figure 11A, right panel). Taken together, these results suggest that *IRTA1* is expressed in a lymphoid subpopulation present in spleen and lymph nodes and presumably represented by resting B cells. 25

To further investigate the phenotype and tissue distribution of the cells expressing *IRTA1*, we performed 30 *in situ* hybridization on human tonsillar tissue using a *IRTA1* antisense cDNA probe (Figure 11B). Serial sections were processed for *in situ* hybridization with a control

-70-

sense cDNA probe (Panel # 1 in Figure 11B), an antisense cDNA probe (Panel # 2) and hematoxylin and eosin (H&E) staining (Panel # 3) to outline the architecture of the lymphoid tissue. The *IRTA1* hybridization signal was  
5 excluded from the germinal center and the mantle zone of the follicles and was characteristically concentrated in the perifollicular zone with infiltrations in the intraepithelial region (Figures 11B-2, 11B-4). In this region,  
only B cells were positive as documented by staining with  
10 B cell specific markers (IgD, not shown), and by immunohistochemical analysis with anti-*IRTA1* and anti-B (CD20, PAX5), anti-T (CD3), and anti-monocyte (CD68) antibodies (not shown; G. Cattoretti et al., manuscript in preparation). This perifollicular area is the "marginal  
15 zone" equivalent of the tonsil, representing a functionally distinct B-cell compartment that contains mostly memory B-cells and monocytid B-cells (de Wolf-Peeters et al., 1997). Together with the Northern blot analysis of normal tissues and cell lines, these results indicate that *IRTA1*  
20 is expressed in a subpopulation of resting mature B-cells topographically located in the perifollicular and intraepithelial region, sites rich in memory B cells.

In the case of *IRTA2*, Northern blot analysis detected all  
25 alternatively spliced species in human lymph node, spleen, bone marrow and small intestine mRNA, with relative preponderance of the *IRTA2a* isoform (Figure 12A, left panel). Among the hematopoietic cell lines of lymphoid and non-lymphoid origin tested, *IRTA2* expression was  
30 restricted to B-cell lines with an immunoblastic, post-germinal center phenotype (Figure 12A, right panel). Similarly to *IRTA1*, it was absent from cell lines derived

-71-

from pre-B cells, germinal center centroblasts, plasma cells, T-cells, erythroid cells and myeloid cells (Figure 12A, right panel).

5     *In situ* hybridization analysis of human tonsillar tissue, using the *IRTA2c* cDNA as a probe, was consistent with the results of the Northern blot analysis. The *IRTA2* mRNA was largely excluded from the mantle zone of the germinal center, with the exception of a few positive cells  
10    (Figures 12B-2, 12B4). Within the germinal center, the dark zone, represented by centroblasts, appeared negative for *IRTA2*, while the light zone, rich in centrocytes, was strongly positive (Figures 12B-2, 12B-4). Finally, *IRTA2* mRNA was detected in the "marginal zone" equivalent region  
15    outside germinal center follicles and in the intraepithelial and interfollicular regions of the tonsil. This pattern is consistent with specificity of *IRTA2* for centrocytes and post-germinal center B cells. Comparing their expression patterns, we conclude that both are  
20    specific for mature B cells, but *IRTA2* has a broader pattern of expression that includes centrocytes and interfollicular B cells, while *IRTA1* is restricted to marginal zone B cells, most likely memory cells.

25    **Genomic organization of the *IRTA1* and *IRTA2* genes**

To understand the consequences of 1q21 abnormalities on *IRTA1* and *IRTA2* gene structure and expression, we first determined the organization of their genomic loci. The *IRTA1* gene contains 11 exons with a total genomic size of 24.5 kb (Figure 13). The *IRTA2* locus was found to span a genomic region of approximately 40 kb (Figure 13). The three *IRTA2* alternatively spliced products share their

-72-

first 8 exons, at which point *IRTA2b* does not utilize the next splicing site, and terminates by entering its 3'UTR region. *IRTA2a* and *2c* isoforms splice into exon 9, with *IRTA2a* entering into its 3'UTR after exon 11 and *IRTA2c* splicing into exon 12 and extending until exon 18 (Figure 5 13).

Based on sequencing data, we determined that the *IRTA1* and *IRTA2* genes are located 21 kb distant from each other, 10 juxtaposed in the same transcriptional orientation (Figure 13) that extends from the telomere (5') towards the centromere (3'). At the 1q21 locus, they are tightly linked to each other as well as to three additional genes we recently cloned through their homology to the *IRTAs* 15 (I.M., manuscript in preparation). All five genes are contiguous, covering a ~300 kb region at 1q21. This region is located at the interval between previously reported 1q21 breakpoints. Based on the distance between genomic clones harboring the respective genes on the 20 Whitehead Institute Radiation Hybrid map, the *IRTA1-2* locus is estimated to lie approximately 0.8 Mb away from the *MUC1* locus towards the telomere (N.P., unpublished data; Dyomin et al., 2000; Gilles et al., 2000) and less than or equal to 7 Mb away from the *FCCRIIB* locus towards 25 the centromere (N.P., unpublished data).

The *t(1;14)(q21;q32)* translocation generates an *IRTA1/Ca<sub>1</sub>* fusion protein in the FR4 myeloma cell line

30 Comparative restriction and nucleotide sequence analysis of germline versus rearranged sequences from the *Ca<sub>1</sub>* and *IRTA1* loci showed that the translocation had fused

-73-

sequences within intron 2 of the *IRTA1* gene to the intronic sequences between the CH3 and the transmembrane exon of *Ca<sub>1</sub>* in the same transcriptional orientation (Figure 14A). This suggested that, if *IRTA1* sequences were expressed in the translocated locus, the intact donor site at the 3' border of the *IRTA1* exon and the intact acceptor site at the 5' of *Ca<sub>1</sub>* could be used to generate a fusion *IRTA1/Ca<sub>1</sub>* mRNA, and possibly a *IRTA1/Ca<sub>1</sub>* fusion protein.

In order to test this prediction, we analyzed *IRTA1* mRNA expression in FR4 by Northern blot analysis using an *IRTA1* cDNA probe derived from exon 1 (Figure 14A). This probe detected a 0.8 kb message in FR4 that was absent from other B-cell lines, and was shorter than the normal 2.5 kb message detectable in ER/EB cells (Figure 14B). We cloned this transcript by RT-PCR of FR4 mRNA using primers derived from sequences at the 5' border of *IRTA1* exon 1 and the 3' border of the *Ca<sub>1</sub>* cytoplasmic exon (Figure 14A). An RT-PCR product was obtained from FR4, but not from the DAKIKI cell line expressing wild-type surface IgA, or other cell lines lacking a t(1;14) translocation (data not shown). Direct sequencing analysis of the PCR product indicated that splicing had precisely linked *IRTA1* and *Ca<sub>1</sub>* at canonical splicing sites and determined that the fusion transcript was 820 bp long.

Analysis of the predicted protein product indicated that the *IRTA1/Ca<sub>1</sub>* splicing had resulted in a fusion between the *IRTA1* signal peptide and first two extracellular aminoacids, with the 32-amino acid long extracellular spacer, transmembrane domain and cytoplasmic tail of the

-74-

- membrane IgA<sub>1</sub> (mIgA<sub>1</sub>) receptor (Figure 14C). To assay for the expression of this fusion protein in FR4 protein extracts, we used an antibody directed against extracellular aminoacid residues specific for the transmembrane isoform of Ca<sub>1</sub> (Yu et al., 1990) for immunoprecipitation, followed by Western blotting. Our results demonstrated that FR4 cells, but not a control cell line (DAKIKI) expressing wild-type surface IgA, express a 9.8kDa protein consistent with the predicted size of IRTA1/Ca<sub>1</sub> fusion protein (Figure 14D). These results show that the translocated allele encodes a fusion protein, composed of the signal peptide and first two extracellular residues of IRTA1 (17 aminoacids) fused to the Ca<sub>1</sub> encoded transmembrane and cytoplasmic domains (71 aminoacids). In contrast to IRTA1/Ca<sub>1</sub>, overexpression on der(14), no expression was detected in FR4 for the reciprocal Ca<sub>1</sub>/IRTA1 transcript or for the intact IRTA2 gene on der(1).
- With the exception of FR4, IRTA1 mRNA expression was not detected in any other myeloma or lymphoma cell line, regardless of the status of its chromosomal band 1q21 (data not shown). Thus, the IRTA1/Ca fusion represents a rare event in 1q21 aberrations.
- Frequent deregulation of IRTA2 expression in cell lines carrying 1q21 abnormalities**
- In order to establish the physical relationship between other 1q21 breakpoints and the IRTA1/2 locus, we performed FISH analysis with the PAC 49A16 on our panel of BL and MM cell lines. Among ten BL cell lines analyzed, seven with

-75-

dup(1)(q21q32) and three with 1q21 translocations (AS283A, BL104, BL136), we detected three signals corresponding to the *IRTA1/IRTA2* locus in seven of the former and two of the latter, consistent with dup(1)(q21q32) in the first 5 case and dup(1)(q21q32) followed by a translocation breakpoint at 1q21 in the second. (Table 1). FISH analysis of AS283A and BL136, using probes spanning the *IRTA* locus and with neighboring genomic clones, placed the 10 breakpoint of the derivative chromosomes outside the *IRTA* locus in both cell lines, at a distance of >800 kb towards the centromere in AS283A and >800 kb towards the telomere in BL136 (N.P, unpublished results). Consistent with this 15 finding, analysis of 30 cases of MM primary tumors by interphase FISH with the 300-kb YAC 23GC4 (Figure 13), showed that 15 cases (50% of total analyzed) had more than 20 two interphase FISH signals (data not shown), while double color FISH with two PAC clones flanking the YAC centromeric and telomeric borders detected no split of these two probes in any of the cases. These results indicate that, with the exception of FR4, the breakpoints 25 of 1q21 aberrations in BL or MM are not within or in close proximity to the genomic region defined by *IRTA1* and *IRTA2*. However, the consistent outcome of either dup(1)(q21q32) (see Table 1) or dup(1)(q21q32) followed by unbalanced translocations (AS283A, BL136, XG2, XG7 in Table 1) is partial trisomy or tetrasomy of the region of 1q21 containing the *IRTA* genes.

-76-

Table 1. Summary of karyotypic and FISH data on IRTA1/IRTA2 locus

Tumour type	Cytogenetics	PAC 49A16	Copy number of IRTA2 locus by FISH	IRTA2 mRNA expression
Burkitt Lymphoma AS283A	der(4) t(1;4)(q21;q35)	der(4), normal 1	3	+++++
MCL16	dup1q21	dup1q21	3	+++
CA46	dup1q21	dup1q21	3	+++
PA682	dup1q21	dup1q21	3	++
BR1gA	dup1q21	dup1q21	3	++
BL32	dup1q21	dup1q21	3	-
BL92	dup1q21	dup1q21	3	++
BL103	inv dup1q21	dup1q21	3	+
BL104	t(1;3)(q21;p25)	der(1)	2	+
BL136	der(1)(qpter1q21::q21)	der(1)	3	++
Multiple Myeloma				
XG2	der(1) t(1;?) (q21;?) der 19 t (1;19) (q12;?)	der(1), normal 1 der(19)	3	++++
XG7	der(9) t (1;9) (q12;?) der(19) t (1;19) (q12;?) der(1) t(1;?) (q21;?) x2	der(9) der(19) der(1) x2	4	-

-77-

We then investigated whether these aberrations had an effect on *IRTA2* mRNA expression. To this end, we used a cDNA probe corresponding to the *IRTA2* 5' untranslated region to screen a Northern blot with a panel of B-NHL and MM cell lines lacking or displaying 1q21 chromosomal abnormalities. The results show that most (ten out of twelve) BL lines with normal 1q21 chromosomes essentially lack *IRTA2* expression, consistent with the fact that BL derive from GC centroblasts which normally lack *IRTA2* expression (Figure 15A, left panel). In contrast, most BL lines carrying 1q21 abnormalities (ten out of twelve) clearly display *IRTA2* mRNA upregulation (Figure 15A, right panel), ranging from 2 to 50 fold over baseline levels detected in BL with normal 1q21. Among myeloma cell lines, *IRTA2* was overexpressed in one out of three lines displaying 1q21 abnormalities (XG2), while it was expressed in none out of seven with normal 1q21 (Figure 15B).

These results show a strong correlation between the presence of 1q21 chromosomal aberrations and deregulation of *IRTA2* mRNA expression in BL and suggest that trisomies of the *IRTA2* locus may deregulate its expression in this lymphoma subtype (see Discussion).

25

#### Discussion

Efforts described herein to identify genes involved in chromosomal aberrations affecting band 1q21 in Multiple Myeloma and B cell lymphoma, led to the discovery of *IRTA1* and *IRTA2*, two founding members of a novel subfamily of related receptors within the immunoreceptor family; full length nucleic acid sequences encoding *IRTA1* and *IRTA2*

-78-

proteins are provided herein, as are the amino acid sequences of the encoded IRTA1 and IRTA2 proteins. Subsequently three additional genes of members of this subfamily of related receptors were isolated, IRTA3, 5 IRTA4, and IRTA5, the full length nucleic acid sequences of which are provided herein, as are the amino acid sequences of the encoded IRTA3, IRTA4, and IRTA5 proteins. These results have implications for the normal biology of, B cells as well as for the role of 1q21 aberrations in 10 lymphomagenesis.

**IRTA1 and IRTA2 are founding members of a new subfamily within the Ig superfamily**

15 Several features shared between the two IRTA genes and their encoded proteins suggest that they form a new subfamily within the immunoreceptor superfamily. First, they share a higher degree of homology with each other in their extracellular domains than with other superfamily members both in their mRNA (68% identity) and protein (47% identity) sequence. Second, they share homology in their cytoplasmic domains, marked by the presence of ITAM-like and ITIM signaling motifs in the context of homologous aminoacid sequences. Third, IRTA1 and IRTA2 belong to a 20 larger subfamily of five genes displaying higher intrafamily homology and tight clustering within a ~300 kb region at 1q21 (I.M. et al., manuscript in preparation). Their genomic organization suggests that a common ancestral gene may have given rise to this subfamily, by a 25 process of duplication and sequence divergence, similar to the mechanism proposed for the Fc receptor family (Qiu et al., 1990).

30 In their extracellular domain, the IRTA proteins are closely related to the Fc receptor subfamily based on the high degree of aminoacid homology shared especially with the high affinity FCGRI receptor (37-45% aminoacid

-79-

identity). A common evolutionary origin with Fc receptors is also suggested by the position of the *IRTA* family locus in the interval between the *FCGRI* locus on 1q21 and the *FCERI* and *FCGRII-III* loci on 1q21-q23. Finally, the *IRTA* 5 and *FCR* genes share a similar exon/intron organization of the gene portion that encodes their signal peptide, in particular the two 5' leader exons with the sequences encoding the signal peptidase site located within the second 21-bp exon.

10 Based on their cytoplasmic ITIM-like motifs, the *IRTA*proteins can be considered members of the Inhibitory Receptor Superfamily (IRS), a group of receptors that block activation of many cell types in the immune system 15 (Lanier, 1998). Such members include *FCGRIIB* and *CD22* in the human (DeLisser et al., 1994) and *PIR-B* in the mouse (Kubagawa et al., 1997). Analogous to IRS members, the ITIM of *IRTA1* and *IRTA2* are encoded by individual exons. A feature that many IRS members share is the existence of 20 corresponding activating receptor isoforms whose cytoplasmic domains are devoid of ITIM (reviewed in Ravetch and Lanier, 1998). It is possible that the secreted isoform of *IRTA2*, which lacks ITIM-like motifs, fulfills an analogous role by counteracting the effect of 25 the transmembrane isoform.

Significant homology in the sequence and overall organization of their extracellular portion is shared among the *IRTA1* and *IRTA2* proteins and the Cell Adhesion 30 Molecule (CAM) subfamily members *PECAM1*, *CD22* and *BGP1*. In addition, the ability of *IRTA2* to generate three protein isoforms with distinct subcellular localization (a transmembrane, a GPI-linked or a secreted protein) by differential splicing is shared by *NCAM*, another member of 35 the CAM subfamily (Dickson et al., 1987; Gower et al., 1988). Thus, the *IRTA* family is also related to the CAM family, as has been previously suggested for a member of the Fc receptor family (murine *FCGRII*) because of its

-80-

homology to PECAM1 (CAM, IRS family) (Daeron, 1991; Newman et al., 1990; Stockinger et al., 1990).

In conclusion, the IRTA family may represent an  
5 intersection among the Fc, IRS and CAM families, combining features from all three. Accordingly, IRTA proteins may have a role in the regulation of signal transduction during an immune response (like Fc receptors), intercellular communication (like members of the IRS and  
10 CAM families) and cell migration (like CAM family members) (DeLisser et al., 1994; Ravetch and Lanier, 2000). Initial experiments indicate that IRTA1 can weakly bind heat aggregated IgA, while IRTA2c can specifically bind heat aggregated human serum IgG (with higher affinity for  
15 IgG<sub>1</sub> and IgG<sub>2</sub>), but not monomeric human IgG, IgA, IgM and IgE (data not shown). These initial data lend support to a functional relationship between the IRTA and the Fc receptor families, but do not exclude functions dependent on other ligands for the IRTA proteins.

20

**Differential pattern of expression of IRTA genes in mature B cells**

The IRTA genes display a specific pattern of expression in  
25 various normal B cell compartments. IRTA1 is topographically restricted to B cells within the perifollicular region, which was originally named marginal zone in the spleen, but is also detectable in most lymphoid organs (de Wolf-Peeters et al., 1997). The *in situ* hybridization data presented here have been confirmed  
30 by immunohistochemical analysis using anti-IRTA1 antibodies which show that the IRTA1 protein is selectively expressed in marginal zone B cells, and, among NHL, in marginal zone lymphoma, the tumors deriving from  
35 these cells (G. Cattoretti et al., manuscript in preparation). On the other hand, IRTA2 has a broader pattern of expression that includes GC centrocytes, as well as a broad spectrum of perifollicular cells, which

-81-

may include immunoblasts and memory cells. Initial data suggest that the pattern of expression of *IRTA3* is analogous to *IRTA2*, while *IRTA4* and *IRTA5* are selectively expressed in mantle zone B cells (I. Miller et al., 5 manuscript in preparation), the pre-GC compartment of mature B cells (MacLennan, I. C., 1994). This topographic restriction of *IRTA* gene expression in lymphoid organs suggests that the *IRTA* molecules may play a role in the migration or activity of various B cell subpopulations in .10 specific functional B cell compartments. In addition, *IRTA* expression should be useful for the differential diagnosis of NHL subtypes deriving from various B cell compartments, particularly *IRTA1* in the diagnosis of marginal zone lymphoma.

15

#### *IRTA1* locus and 1q21 abnormalities in MM

In the FR4 cell line, the consequence of the t(1;14) translocation is the formation of an *IRTA1/Cα<sub>1</sub>* fusion gene. Despite the fact that this gene is driven by the *IRTA1* promoter region, which is normally silent in plasma cells, 20 its expression is high in FR4, presumably due to the influence of the *Cα<sub>1</sub>* 3' LCR, which is retained downstream of the *Cα<sub>1</sub>* locus. The fusion gene encodes a *IRTA1/Cα<sub>1</sub>* fusion protein which contains only the signal peptide and 25 first two amino acids of *IRTA1* linked to the surface IgA receptor. The latter has been almost completely deprived of its extracellular domain, but retains all its transmembrane and intracellular domains. This structure indicates that the *IRTA1/Cα<sub>1</sub>* fusion protein, though 30 probably unable to bind any ligand, may retain the potential for dimerization and signaling. In particular, the membrane (m) IgA-derived extracellular portion contains a cysteine residue, which can be involved in disulphide bonds between two α-chains or between α-chains 35 and associated proteins, such as the auxilliary surface receptor CD19 (Leduc et al., 1997). The fusion protein also carries the intact, 14 amino acid mIgA cytoplasmic

-82-

domain, which is highly conserved in evolution (Reth, 1992) and may play an essential role in the proliferation, survival and differentiation of mature B-cells, analogous to the role of mIgG and mIgE (Kaisho et al., 1997). Thus, 5 the emergence of the IRTA1/Ca<sub>1</sub> protein in FR4 may have provided the cells with a proliferative and survival advantage during tumor development through ligand (antigen)-independent activation of the BCR pathway. This fusion event however, appears to be rare in B-cell 10 malignancy, since so far we were able to detect it only in FR4 cells.

#### IRTA2 locus and 1q21 abnormalities in MM and BL

Abnormal expression of IRTA2 is a frequent consequence of 15 1q21 abnormalities. Although this gene is not expressed normally either in centroblasts, the presumed normal counterparts of BL (Kuppers et al., 1999), or in BL with normal 1q21, its levels are upregulated on average by 10-fold in BL cell lines with 1q21 abnormalities. This 20 deregulation appears to be specific for IRTA2 since all the other 4 IRTA genes present within 300 kb on 1q21 are either not expressed in BL (IRTA1), or their pattern of expression does not correlate with the presence of 1q21 25 abnormalities (IRTA3, 4, 5, not shown). The mechanism by which this deregulation occurs is difficult to ascertain in the absence of structural lesions within or adjacent to the IRTA2 gene. Since the heterogeneous aberrations that affect 1q21 all cause an excess copy number of the IRTA 30 locus, it is possible that this may lead to regulatory disturbances, as is the case for low level amplification of BCL2 in FL lacking (14;18) translocations (Monni et al., 1997), REL in diffuse large cell lymphoma (Houldsworth et al., 1996; Rao et al., 1998) and 35 deregulation of Cyclin D1 in some MM cases with trisomy 11 (Pruneri et al., 2000). On the other hand, 1q21 abnormalities, including translocations and duplications, change the genomic context of the IRTA locus and may lead

-83-

to deregulation of *IRTA2* by distant cis-acting enhancer chromatin organizing elements acting on its promoter as is the case for *MYC* in endemic BL (Pellicci et al., 1986) and MM (Shou et al., 2000) and for *CCND1* in mantle cell lymphoma (Bosch et al., 1994; Swerdlow et al., 1995) and MM (Pruneri et al., 2000).

The biological consequences of deregulated *IRTA2* expression are difficult to predict at this stage. The observation that *IRTA2* has homology with CAM adhesion receptors, together with its specific distribution in the light zone of the GC suggest that its ectopic expression in centroblasts may cause a disruption in the GC development and architecture. On the other hand, our initial observations that *IRTA2* can bind IgG immune complexes comparably to bona fide Fc receptors suggest that its inappropriate expression may perturb the dynamics of cell surface regulation of B cell immunological responses, possibly leading to clonal expansion.

Deregulated expression of *FCGR2B* as a result of the t(1;14)(q21;q32) in follicular lymphoma has been proposed to contribute to lymphomagenesis in this tumor type (Callanan et al., 2000), by a mechanism involving escape by tumor cells of anti-tumor immune surveillance through their Fc binding and inactivation of tumor specific IgG. Similar evasion mechanisms have been observed in cells infected by Fc-encoding herpesviruses (Dubin et al., 1991). The role of *IRTA2* deregulation needs to be tested in "gain of function" transgenic mice constitutively expressing *IRTA2* in the GC.

-84-

References for Second Series of Experiments

- Anand, R., Riley, J. H., Butler, R., Smith, J. C., and Markham, A. F. (1990). A 3.5 genome equivalent multi access YAC library: construction, characterisation, 5 screening and storage. *Nucleic Acids Res* 18, 1951-6.
- Avet-Loiseau, H., Andree-Ashley, L. E., Moore, D., 2nd, Mellerin, M. P., Feusner, J., Bataille, R., and Pallavicini, M. G. (1997). Molecular cytogenetic 10 abnormalities in multiple myeloma and plasma cell leukemia measured using comparative genomic hybridization. *Genes Chromosomes Cancer* 19, 124-33.
- Bakhshi, A., Jensen, J. P., Goldman, P., Wright, J. J., 15 McBride, O. W., Epstein, A. L., and Korsmeyer, S. J. (1985). Cloning the chromosomal breakpoint of t(14;18) human lymphomas: clustering around JH on chromosome 14 and near a transcriptional unit on 18. *Cell* 41, 899-906.
- Berger, R., Bernheim, A. (1985). Cytogenetics of Burkitt's 20 lymphoma-leukaemia: a review. *IARC Sci Publ* 60, 65-80.
- Bergsagel, P. L., Chesi, M., Nardini, E., Brents, L. A., Kirby, S. L., and Kuehl, W. M. (1996). Promiscuous 25 translocations into immunoglobulin heavy chain switch regions in multiple myeloma. *Proc Natl Acad Sci U S A* 93, 13931-6.
- Bosch, F., Jares, P., Campo, E., Lopez-Guillermo, A., 30 Piris, M. A., Villamor, N., Tassies, D., Jaffe, E. S., Montserrat, E., Rozman, C. et al. (1994). PRAD-1/cyclin D1 gene overexpression in chronic lymphoproliferative disorders: a highly specific marker of mantle cell lymphoma. *Blood* 84, 2726-32.

-85-

Callanan, M. B., Le Baccon, P., Mossuz, P., Duley, S.,  
Bastard, C., Hamoudi, R., Dyer, M. J., Klobeck, G.,  
Rimokh, R., Sotto, J. J., and Leroux, D. (2000). The IgG  
Fc receptor, FcgammaRIIB, is a target for deregulation by  
5 chromosomal translocation in malignant lymphoma. Proc Natl  
Acad Sci U S A 97, 309-14.

Cambier, J. C. (1995). Antigen and Fc receptor signaling.  
The awesome power of the immunoreceptor tyrosine-based  
10 activation motif (ITAM). J Immunol 155, 3281-5.

Chesi, M., Bergsagel, P. L., Brents, L. A., Smith, C. M.;  
Gerhard, D. S., and Kuehl, W. M. (1996). Dysregulation of  
cyclin D1 by translocation into an IgH gamma switch region  
15 in two multiple myeloma cell lines [see comments]. Blood  
88, 674-81.

Chesi, M., Bergsagel, P. L., Shonukan, O. O., Martelli, M.  
L., Brents, L. A., Chen, T., Schrock, E., Ried, T., and  
20 Kuehl, W. M. (1998). Frequent dysregulation of the c-maf  
proto-oncogene at 16q23 by translocation to an Ig locus in  
multiple myeloma. Blood 91, 4457-63.

Chesi, M., Nardini, E., Brents, L. A., Schrock, E., Ried,  
25 T., Kuehl, W. M., and Bergsagel, P. L. (1997). Frequent  
translocation t(4;14) (p16.3;q32.3) in multiple myeloma is  
associated with increased expression and activating  
mutations of fibroblast growth factor receptor 3. Nat  
Genet 16, 260-4.

30

Church, D. M., Stotler, C. J., Rutter, J. L., Murrell, J.  
R., Trofatter, J. A., and Buckler, A. J. (1994). Isolation

-86-

of genes from complex sources of mammalian genomic DNA using exon amplification. *Nat Genet* 6, 98-105.

Cigudosa, J. C., Parsa, N. Z., Louie, D. C., Filippa, D.

5 A., Jhanwar, S. C., Johansson, B., Mitelman, F., and Chaganti, R. S. (1999). Cytogenetic analysis of 363 consecutively ascertained diffuse large B-cell lymphomas. *Genes Chromosomes Cancer* 25, 123-33.

10 Daeron, M. (1991). Fc receptors, or the elective affinities of adhesion molecules. *Immunol Lett* 27, 175-81.

Dalla-Favera, R., Bregni, M., Erikson, J., Patterson, D., Gallo, R.C., and Croce, C.M.: The human c-myc onc-gene is 15 located on the region of chromosome 8 which is translocated in Burkitt lymphoma cells. *Proc. Nat. Acad. Sci. USA* 79:7824-7827, 1982.

Dalla-Favera, R., Martinotti, S., Gallo, R. C., Erikson, 20 J., and Croce, C. M. (1983). Translocation and rearrangements of the c-myc oncogene locus in human undifferentiated B-cell lymphomas. *Science* 219, 963-7.

de Wolf-Peeters, C., Pittaluga, S., Dierlamm, J., 25 Wlodarska, I., and Van Den Berghe, H. (1997). Marginal zone B-cell lymphomas including mucosa-associated lymphoid tissue type lymphoma (MALT), monocyteid B-cell lymphoma and splenic marginal zone cell lymphoma and their relation to the reactive marginal zone. *Leuk Lymphoma* 26, 467-78.

-87-

DeLisser, H. M., Newman, P. J., and Albelda, S. M. (1994). Molecular and functional aspects of PECAM-1/CD31. *Immunol Today* 15, 490-5.

5 Dickson, G., Gower, H. J., Barton, C. H., Prentice, H. M., Elsom, V. L., Moore, S. E., Cox, R. D., Quinn, C., Putt, W., and Walsh, F. S. (1987). Human muscle neural cell adhesion molecule (N-CAM): identification of a muscle-specific sequence in the extracellular domain. *Cell* 50, 10 1119-30.

Dierlamm, J., Pittaluga, S., Wlodarska, I., Stul, M., Thomas, J., Boogaerts, M., Michaux, L., Driessens, A., Mecucci, C., Cassiman, J. J., and et al. (1996). Marginal 15 zone B-cell lymphomas of different sites share similar cytogenetic and morphologic features [see comments]. *Blood* 87, 299-307.

Dracopoli, C. N., Haines, J. L., Korf, B. R., Morton, C. 20 C., Seidman, C. E., Seidman, J.G., Smith, D. R. (1997). Current Protocols in Human Genetics (New York: Wiley & Sons)

Dubin, G., Socolof, E., Frank, I., Friedman, H. M. (1991). 25 Herpes simplex virus type 1 Fc receptor protects infected cells from antibody-dependent cellular cytotoxicity. *Journal of Virology* 65, 7046-50.

Dyomin, V. G., Palanisamy, N., Lloyd, K. O., Dyomina, K., 30 Jhanwar, S. C., Houldsworth, J., and Chaganti, R. S. (2000). MUC1 is activated in a B-cell lymphoma by the

-88-

t(1;14) (q21;q32) translocation and is rearranged and amplified in B-cell lymphoma subsets. Blood 95, 2666-71.

Dyomin, V.G., Rao, P.H., Dalla-Favera, R., Chaganti,  
5 R.S.K. (1997). *BCL8*, a novel gene involved in translocations affecting band 15q11-13 in diffuse large-cell lymphoma. Proc Natl Acad Sci USA 94, 5728-32.

Eton, O., Scheinberg, D. A., and Houghton, A. N. (1989).  
10 Establishment and characterization of two human myeloma cell lines secreting kappa light chains. Leukemia 3, 729-35.

Ferguson, M. A., and Williams, A. F. (1988). Cell-surface anchoring of proteins via glycosyl-phosphatidylinositol structures. Annu Rev Biochem 57, 285-320.

Frank, D., Mendelsohn, C. L., Ciccone, E., Svensson, K., Ohlsson, R., and Tycko, B. (1999). A novel pleckstrin homology-related gene family defined by Ipl/Tssc3, TDAG51, and Tih1: tissue-specific expression, chromosomal location, and parental imprinting. Mamm Genome 10, 1150-1159.

25 Gaidano, G., and Dalla-Favera, R. (1997). Molecular Biology of Lymphomas. In: Principles and Practice of Oncology, Fifth Ed, DeVita, VT, Hellman, S., Rosenberg SA (eds) JB Lippincott Co (publ.), 2131-2145.

30 Gilles, F., Goy, A., Remache, Y., Shue, P., and Zelenetz, A. D. (2000). MUC1 dysregulation as the consequence of a

-89-

tt(1;14) (q21;q32) translocation in an extranodal lymphoma.  
Blood 95, 2930-2936.

- Gower, H. J., Barton, C. H., Elsom, V. L., Thompson, J.,  
5 Moore, S. E., Dickson, G., and Walsh, F. S. (1988). Alternative splicing generates a secreted form of N-CAM in muscle and brain. Cell 55, 955-64.
- Hamilton, M. S., Ball, J., Bromidge, E., Lowe, J., and  
10 Franklin, I. M. (1990). Characterization of new IgG lambda myeloma plasma cell line (EJM): a further tool in the investigation of the biology of multiple myeloma. Br J Haematol 75, 378-84.
- Houldsworth, J., Mathew, S., Rao, P. H., Dyomina, K., Louie, D. C., Parsa, N., Offit, K., Chaganti, R. S. (1996). REL proto-oncogene is frequently amplified in extranodal diffuse large cell lymphoma. Blood 87, 25-9.
- Iida, S., Rao, P. H., Butler, M., Corradini, P., Boccadoro, M., Klein, B., Chaganti, R. S., and Dalla-Favera, R. (1997). Dereulation of MUM1/IRF4 by chromosomal translocation in multiple myeloma. Nat Genet 17, 226-30.
- 25 Jackson, N., Lowe, J., Ball, J., Bromidge, E., Ling, N. R., Larkins, S., Griffith, M. J., and Franklin, I. M. (1989). Two new IgA1-kappa plasma cell leukaemia cell lines (JJN-1 & JJN-2) which proliferate in response to B cell stimulatory factor 2. Clin Exp Immunol 75, 93-9.

-90-

- Jernberg, H., Zech, L., and Nilsson, K. (1987). Cytogenetic studies on human myeloma cell lines. *Int J Cancer* 40, 811-7.
- 5 Juliusson, G., Oscier, D. G., Fitchett, M., Ross, F. M., Stockdill, G., Mackie, M. J., Parker, A. C., Castoldi, G. L., Guneo, A., Knuutila, S., and et al. (1990). Prognostic subgroups in B-cell chronic lymphocytic leukemia defined by specific chromosomal abnormalities. *N Engl J Med* 323, 10 720-4.
- Kaisho, T., Schwenk, F., and Rajewsky, K. (1997). The roles of gamma 1 heavy chain membrane expression and cytoplasmic tail in IgG1 responses. *Science* 276, 412-5.
- 15 Kempkes, B., Spitkovsky, D., Jansen-Durr, P., Ellwart, J. W., Kremmer, E., Delecluse, H. J., Rottenberger, C., Bornkamm, G. W., and Hammerschmidt, W. (1995). B-cell proliferation and induction of early G1-regulating 20 proteins by Epstein-Barr virus mutants conditional for EBNA2. *Embo J* 14, 88-96.
- Kornblau, S. M., Goodacre, A., Cabanillas, F. (1991). Chromosomal abnormalities in adult non-endemic Burkitt's 25 lymphoma and leukemia: 22 new reports and a review of 148 cases from the literature. *Hematol Oncol* 9, 63-78.
- Kubagawa, H., Burrows, P. D., and Cooper, M. D. (1997). A novel pair of immunoglobulin-like receptors expressed by B 30 cells and myeloid cells [see comments]. *Proc Natl Acad Sci U S A* 94, 5261-6.

-91-

- Kuppers, R., Klein, U., Hansmann, M. L., and Rajewsky, K. (1999). Cellular origin of human B-cell lymphomas. *N Engl J Med* 341, 1520-9.
- 5 Lanier, L. L. (1998). NK cell receptors. *Annu Rev Immunol* 16, 359-93.
- Leduc, I., Drouet, M., Bodinier, M. C., Helal, A., and Cogne, M. (1997). Membrane isoforms of human 10 immunoglobulins of the A1 and A2 isotypes: structural and functional study. *Immunology* 90, 330-6.
- MacLennan, I. C. (1994). Germinal Centers. *Annu Rev Immunol* 12, 117-39.
- 15 Magrath, I. T., Pizzo, P. A., Whang-Peng, J., Douglass, E. C., Alabaster, O., Gerber, P., Freeman, C. B., and Novikovs, L. (1980). Characterization of lymphoma-derived cell lines: comparison of cell lines positive and negative 20 for Epstein-Barr virus nuclear antigen. I. Physical, cytogenetic, and growth characteristics. *J Natl Cancer Inst* 64, 465-76.
- Monni, O., Joensuu, H., Franssila, K., Klefstrom, J., Alitalo, K., and Knuutila, S. (1997). BCL2 overexpression associated with chromosomal amplification in diffuse large 25 B-cell lymphoma. *Blood* 90, 1168-74.
- Neri, A., Barriga, F., Knowles, D. M., Magrath, I. T., and Dalla-Favera, R. (1988). Different regions of the 30 immunoglobulin heavy-chain locus are involved in

-92-

chromosomal translocations in distinct pathogenetic forms of Burkitt lymphoma. Proc Natl Acad Sci U S A 85, 2748-52.

- Neri, A., Chang, C. C., Lombardi, L., Salina, M.,  
5 Corradini, P., Maiolo, A. T., Chaganti, R. S., and Dalla-  
Favera, R. (1991). B cell lymphoma-associated chromosomal  
translocation involves candidate oncogene lyt-10,  
homologous to NF-kappa B p50. Cell 67, 1075-87.
- 10 Newman, P. J., Berndt, M. C., Gorski, J., White, G. C. d.,  
Lyman, S., Paddock, C., and Muller, W. A. (1990). PECAM-1  
(CD31) cloning and relation to adhesion molecules of the  
immunoglobulin gene superfamily. Science 247, 1219-22.
- 15 Offit, K., Louie, D. C., Parsa, N. Z., Roy, P., Leung, D.,  
Lo Coco, F., Zelenetz, A., Dalla-Favera, R., Chaganti, R.  
S. (1995). BCL6 gene rearrangement and other cytogenetic  
abnormalities in diffuse large cell lymphoma. Leuk  
Lymphoma 20, 85-9.
- 20 Pelicci, P. G., Knowles, D. M. d., Magrath, I., and Dalla-  
Favera, R. (1986). Chromosomal breakpoints and structural  
alterations of the c-myc locus differ in endemic and  
sporadic forms of Burkitt lymphoma. Proc Natl Acad Sci U S  
25 A 83, 2984-8.
- Polito, P., Cilia, A. M., Gloghini, A., Cozzi, M., Perin,  
T., De Paoli, P., Gaidano, G., and Carbone, A. (1995).  
High frequency of EBV association with non-random  
abnormalities of the chromosome region 1q21-25 in AIDS-  
30 related Burkitt's lymphoma-derived cell lines. Int J  
Cancer 61, 370-4.

-93-

- Pruneri, G., Fabris, S., Baldini, L., Carboni, N., Zagano, S., Colombi, M. A., Ciceri, G., Lombardi, L., Rocchi, M., Buffa, R., Maiolo, A. T., Neri, A. (2000). Immunohistochemical analysis of cyclin D1 shows deregulated expression in multiple myeloma with the t(11;14). *Am J Pathol* 156, 1505-13.
- Qiu, W. Q., de Bruin, D., Brownstein, B. H., Pearse, R., Ravetch, J. V. (1990). Organization of the human and mouse low-affinity Fc gamma R genes: duplication and recombination. *Science* 248, 732-5.
- Rao, P. H., Houldsworth, J., Dyomina, K., Parsa, N. Z., Cigudosa, J. C., Louie, D. C., Popplewell, L., Offit, K., Jhanwar, S. C., and Chaganti, R. S. (1998). Chromosomal and gene amplification in diffuse large B-cell lymphoma. *Blood* 92, 234-40.
- Rao, P. H., Murty, V. V., Gaidano, G., Hauptchein, R., Dalla-Favera, R., and Chaganti, R. S. (1993). Subregional localization of 20 single-copy loci to chromosome 6 by fluorescence *in situ* hybridization. *Genomics* 16, 426-30.
- Ravetch, J. V., and Lanier, L. L. (2000). Immune inhibitory receptors [In Process Citation]. *Science* 290, 84-9.
- Reth, M. (1989). Antigen receptor tail clue [letter]. *Nature* 338, 383-4.
- Reth, M. (1992). Antigen receptors on B lymphocytes. *Annu Rev Immunol* 10, 97-121.

-94-

- Richelda, R., Ronchetti, D., Baldini, L., Cro, L., Viggiano, L., Marzella, R., Rocchi, M., Otsuki, T., Lombardi, L., Maiolo, A. T., Neri, A. (1997). A novel chromosomal translocation t(4; 14) (p16.3; q32) in multiple  
5 myeloma involves the fibroblast growth-factor receptor 3 gene [see comments]. *Blood* 90, 4062-70.
- Riley, J., Butler, R., Ogilvie, D., Finniear, R., Jenner, D., Powell, S., Anand, R., Smith, J. C., and Markham, A.  
10 F. (1990). A novel, rapid method for the isolation of terminal sequences from yeast artificial chromosome (YAC) clones. *Nucleic Acids Res* 18, 2887-90.
- Ronchetti, D., Finelli, P., Richelda, R., Baldini, L.,  
15 Rocchi, M., Viggiano, L., Cuneo, A., Bogni, S., Fabris, S., Lombardi, L., Maiolo, A. T., and Neri, A. (1999). Molecular analysis of 11q13 breakpoints in multiple myeloma. *Blood* 93, 1330-7.
- 20 Rosenberg, C. L., Wong, E., Petty, E. M., Bale, A. E., Tsujimoto, Y., Harris, N. L., and Arnold, A. (1991). PRAD1, a candidate BCL1 oncogene: mapping and expression in centrocytic lymphoma. *Proc Natl Acad Sci U S A* 88, 9638-42.
- 25 Sawyer, J. R., Tricot, G., Mattox, S., Jagannath, S., and Barlogie, B. (1998). Jumping translocations of chromosome 1q in multiple myeloma: evidence for a mechanism involving decondensation of pericentromeric heterochromatin. *Blood*  
30 91, 1732-41.

-95-

- Sawyer, J. R., Waldron, J. A., Jagannath, S., Barlogie, B. (1995). Cytogenetic findings in 200 patients with multiple myeloma. *Cancer Genet Cytogenet* 82, 41-9.
- 5 Shou, Y., Martelli, M. L., Gabrea, A., Qi, Y., Brents, L. A., Roschke, A., Dewald, G., Kirsch, I. R., Bergsagel, P. L., and Kuehl, W. M. (2000). Diverse karyotypic abnormalities of the c-myc locus associated with c- myc dysregulation and tumor progression in multiple myeloma.
- 10 Proc Natl Acad Sci U S A 97, 228-33.
- Stockinger, H., Gadd, S. J., Eher, R., Majdic, O., Schreiber, W., Kasinrerk, W., Strass, B., Schnabl, E., and Knapp, W. (1990). Molecular characterization and 15 functional analysis of the leukocyte surface protein CD31. *J Immunol* 145, 3889-97.
- Swerdlow, S. H., Yang, W. I., Zukerberg, L. R., Harris, N. L., Arnold, A., Williams, M. E. (1995). Expression of 20 cyclin D1 protein in centrocytic/mantle cell lymphomas with and without rearrangement of the BCL1/cyclin D1 gene. *Hum Pathol* 26, 999-1004.
- Tagawa, S., Doi, S., Taniwaki, M., Abe, T., Kanayama, Y., 25 Nojima, J., Matsubara, K., and Kitani, T. (1990). Amylase-producing plasmacytoma cell lines, AD3 and FR4, with der(14)t(8;14) and dic(8)t(1;8) established from ascites. *Leukemia* 4, 600-5.
- 30 Taub, R., Kirsch, I., Morton, C., Lenoir, G., Swan, D., Tronick, S., Aaronson, S., Leder, P. (1982). Translocation of the c-myc gene into the immunoglobulin heavy chain

-96-

- locus in human Burkitt lymphoma and murine plasmacytoma cells. Proc Natl Acad Sci U S A 79, 7837-41.
- Thompson, J. D., Higgins, D. G., Gibson, T. J. CLUSTAL W: improving the sensitivity of progressive multiple sequence 5 alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22, 4673-80.
- Tusnady, G. E., Simon, I. (1998). Principles governing 10 amino acid composition of integral membrane proteins: application to topology prediction. J Mol Bio 283, 489-506.
- Unkeless, J. C., and Jin, J. (1997). Inhibitory receptors, 15 ITIM sequences and phosphatases. Curr Opin Immunol 9, 338-43.
- von Heijne, G. (1986). A new method for predicting signal sequence cleavage sites. Nucleic Acids Res 14, 4683-90.
- 20 Whang-Peng, J., Knutson, T., Jaffe, E. S., Steinberg, S. M., Raffeld, M., Zhao, W. P., Duffey, P., Condron, K., Yano, T., Longo, D. L. (1995). Sequential analysis of 43 patients with non-Hodgkin's lymphoma: clinical 25 correlations with cytogenetic, histologic, immunophenotyping, and molecular studies. Blood 85, 203-16.
- Willis, T. G., Zalcberg, I. R., Coignet, L. J., Wlodarska, 30 I., Stul, M., Jadayel, D. M., Bastard, C., Treleaven, J. G., Catovsky, D., Silva, M. L., and Dyer, M. J. (1998). Molecular cloning of translocation tt(1;14)(q21;q32)

-97-

defines a novel gene (BCL9) at chromosome 1q21. Blood 91, 1873-81.

Ye, B. H., Lista, F., Lo Coco, F., Knowles, D. M., Offit, K., Chaganti, R. S., and Dalla-Favera, R. (1993). Alterations of a zinc finger-encoding gene, BCL-6, in diffuse large-cell lymphoma. Science 262, 747-50.

Yu, L. M., Peng, C., Starnes, S. M., Liou, R. S., and Chang, T. W. (1990). Two isoforms of human membrane-bound alpha Ig resulting from alternative mRNA splicing in the membrane segment. J Immunol 145, 3932-6.

Zhang, X. G., Gaillard, J. P., Robillard, N., Lu, Z. Y., Gu, Z. J., Jourdan, M., Boiron, J. M., Bataille, R., and Klein, B. (1994). Reproducible obtaining of human myeloma cell lines as a model for tumor stem cell study in human multiple myeloma. Blood 83, 3654-63.

20 Third Series of Experiments

Chromosome 1q21 is frequently altered by translocations and duplications in several types of B cell malignancy, including multiple myeloma, Burkitt lymphoma, marginal zone lymphomas, and follicular lymphoma. To identify the genes involved in these aberrations, cloned was the chromosomal breakpoint of a t(1;14)(q21;q32) in the myeloma cell line FR4. A 300kb region spanning the breakpoint contains at least five highly related adjacent genes which encode surface receptor molecules that are members of the immunoglobulin gene superfamily, and thus called IRTA (Immunoglobulin Receptor Translocation

-98-

Associated). The various IRTA molecules have from three to nine extracellular immunoglobulin superfamily domains and are related to the Fc gamma receptors. They have transmembrane and cytoplasmic domains containing ITIM-like 5 and ITAM-like (IRTA-1, IRTA-3, IRTA-4) signaling motifs. In situ hybridization experiments show that all IRTA genes are expressed in the B cell lineage with distinct developmental stage-specific patterns: IRTA-1 is expressed in a marginal B cell pattern. IRTA-2 is expressed in 10 centrocytes and more mature B cells. As a result of the translocation in FR4, IRTA-1 is broken and produces a fusion transcript with the immunoglobulin locus. The IRTA-2 gene, normally silent in centroblasts, is overexpressed in multiple myeloma and in Burkitt lymphoma 15 cell lines carrying 1q21 abnormalities. The data here suggests that IRTA genes are novel B cell regulatory molecules that may also have a role in lymphomagenesis.

-99-

What is claimed is:

1. An isolated nucleic acid molecule which encodes immunoglobulin receptor, Immunoglobulin superfamily Receptor Translocation Associated, IRTA, protein.  
5
2. The isolated nucleic acid molecule of claim 1, wherein the IRTA protein is IRTA1 protein comprising the amino acid sequence set forth in Figure 18A (SEQ ID NO:1).  
10
3. The isolated nucleic acid molecule of claim 1, wherein the IRTA protein is IRTA2 protein comprising the amino acid sequence set forth in Figures 18B-1-18B-3 (SEQ ID NO:3).  
15
4. The isolated nucleic acid molecule of claim 1, wherein the IRTA protein is IRTA3 protein comprising the amino acid sequence set forth in Figures 18C-1-18C-2 (SEQ ID NO:5).  
20
5. The isolated nucleic acid molecule of claim 1, wherein the IRTA protein is IRTA4 protein comprising the amino acid sequence set forth in Figures 18D-1-18D-2 (SEQ ID NO: 7).  
25
6. The isolated nucleic acid molecule of claim 1, wherein the IRTA protein is IRTA5 protein comprising the amino acid sequence set forth in Figures 18E-1-18E-2 (SEQ ID NO: 9).  
30

-100-

7. An isolated nucleic acid molecule of claim 1, wherein the nucleic acid molecule is DNA.
8. The isolated DNA molecule of claim 2, wherein the DNA  
5 is cDNA.
9. The isolated DNA molecule of claim 2, wherein the DNA  
is genomic DNA.
- 10 10. The isolated nucleic acid molecule of claim 1,  
wherein the nucleic acid molecule is an RNA molecule.
11. The isolated DNA molecule of claim 2, wherein the DNA  
molecule is cDNA having the nucleotide sequence set  
15 forth in Figure 18A (SEQ ID NO:2).
12. The isolated DNA molecule of claim 2, wherein the DNA  
molecule is cDNA having the nucleotide sequence set  
forth in Figure 18A (SEQ ID NO:4).
- 20 13. The isolated DNA molecule of claim 2, wherein the DNA  
molecule is cDNA having the nucleotide sequence set  
forth in Figure 18A (SEQ ID NO:6).
- 25 14. The isolated DNA molecule of claim 2, wherein the DNA  
molecule is cDNA having the nucleotide sequence set  
forth in Figure 18A (SEQ ID NO:8).
- 30 15. The isolated DNA molecule of claim 2, wherein the DNA  
molecule is cDNA having the nucleotide sequence set  
forth in Figure 18A (SEQ ID NO:10).

-101-

16. The isolated nucleic acid molecule of claim 1, wherein the nucleic acid molecule encodes a human IRTA1 protein.
- 5 17. The isolated nucleic acid molecule of claim 1 operatively linked to a promoter of DNA transcription.
- 10 18. The isolated nucleic acid molecule of claim 17, wherein the promoter comprises a bacterial, yeast, insect, plant or mammalian promoter.
- 15 19. A vector comprising the nucleic acid molecule of claim 17.
20. The vector of claim 19, wherein the vector is a plasmid.
21. A host cell comprising the vector of claim 20.
- 20 22. The host cell of claim 21, wherein the cell is selected from a group consisting of a bacterial cell, a plant cell, and insect cell and a mammalian cell.
- 25 23. An isolated nucleic acid molecule comprising at least 15 contiguous nucleotides capable of specifically hybridizing with a unique sequence included within the sequence of the isolated nucleic acid molecule encoding IRTA1 protein of claim 1.
- 30 24. The isolated nucleic acid molecule of claim 23 labeled with a detectable marker.

-102-

25. The nucleic acid molecule of claim 24, wherein the detectable marker is selected from the group consisting of a radioactive isotope, enzyme, dye,  
5 biotin, a fluorescent label or a chemiluminescent label.
26. A method for detecting a B cell malignancy or a type of B cell malignancy in a sample from a subject  
10 wherein the B cell malignancy comprises a 1q21 chromosomal rearrangement which comprises:  
15 a) obtaining RNA from the sample from the subject;  
b) contacting the RNA of step (a) with a nucleic acid molecule of at least 15 contiguous nucleotides capable of specifically hybridizing with a unique sequence included within the sequence of an isolated RNA encoding human IRTA protein selected from the group consisting of  
20 human IRTA1, IRTA2, IRTA3, IRTA4 and IRTA5, under conditions permitting hybridization of the RNA of step (a) with the nucleic acid molecule capable of specifically hybridizing with a unique sequence included within the sequence of an isolated RNA encoding human IRTA protein,  
25 wherein the nucleic acid molecule is labeled with a detectable marker; and  
c) detecting any hybridization in step (b), wherein detection of hybridization indicates presence of  
30 B cell malignancy or a type of B cell malignancy in the sample.

-103-

27. The method of claim 26, wherein the detectable marker is radioactive isotope, enzyme, dye, biotin, a fluorescent label or a chemiluminescent label.
- 5 28. The method of claim 26, wherein the B cell malignancy is selected from the group consisting of B cell lymphoma, multiple myeloma, Burkitt's lymphoma, marginal zone lymphoma, diffuse large cell lymphoma and follicular lymphoma cells.
- 10 29. The method of claim 28, wherein the B cell lymphoma is Mucosa-Associated-Lymphoid Tissue B cell lymphoma (MALT).
- 15 30. The method of claim 28, wherein the B cell lymphoma is non-Hodgkin's lymphoma.
31. An antisense oligonucleotide having a sequence capable of specifically hybridizing to an mRNA molecule encoding a human ITRA protein so as to prevent overexpression of the mRNA molecule.
- 20 32. The antisense oligonucleotide of claim 31, wherein the ITRA protein selected from the group consisting of human IRTA1, IRTA2, IRTA3, IRTA4 and IRTA5 protein.
- 25 33. A purified IRTA1 protein comprising the amino acid sequence set forth in Figure 18A (SEQ ID NO:1).
- 30 34. The purified IRTA1 protein of claim 33, wherein the IRTA1 protein is human IRTA1.

-104-

35. A purified IRTA2 protein comprising the amino acid sequence set forth in Figures 18B-1-18B-3 (SEQ ID NO:3).
- 5 36. The purified IRTA2 protein of claim 35, wherein the IRTA2 protein is human IRTA2.
- 10 37. A purified IRTA3 protein comprising the amino acid sequence set forth in Figures 18C-1-18C-2 (SEQ ID NO:5).
38. The purified IRTA3 protein of claim 37, wherein the IRTA3 protein is human IRTA3.
- 15 39. A purified IRTA4 protein comprising the amino acid sequence set forth in Figures 18D-1-18D-2 (SEQ ID NO: 7).
- 20 40. The purified IRTA4 protein of claim 39, wherein the IRTA4 protein is human IRTA4.
41. A purified IRTA5 protein comprising the amino acid sequence set forth in Figures 18E-1-18E-2 (SEQ ID NO: 9).
- 25 42. The purified IRTA5 protein of claim 41, wherein the IRTA5 protein is human IRTA5.
- 30 43. An antibody directed to a purified IRTA protein selected from the group consisting of human IRTA1, IRTA2, IRTA3, IRTA4 and IRTA5.

-105-

44. The antibody of claim 43, wherein the IRTA protein is human IRTA protein.
- 5       45. The antibody of claim 43, wherein the antibody is a monoclonal antibody or a polyclonal antibody.
- 10      46. The antibody of claim 43, wherein the monoclonal antibody is a murine monoclonal antibody or a humanized monoclonal antibody.
- 15      47. The antibody of claim 43, wherein the antibody is conjugated to a therapeutic agent, wherein the therapeutic agent is selected from the group consisting of a radioisotope, a toxin, a toxoid, or a chemotherapeutic agent.
- 20      48. A pharmaceutical composition comprising an amount of the antibody of claim 43 effective to bind to cancer cells expressing an IRTA protein selected from the group consisting of human IRTA1, IRTA2, IRTA3, IRTA4 and IRTA5 so as to prevent growth of the cancer cells and a pharmaceutically acceptable carrier.
- 25      49. The pharmaceutical composition of claim 48, wherein the cancer cells are selected from the group consisting of B cell lymphoma, a mantle cell lymphoma multiple myeloma, Burkitt's lymphoma, marginal zone lymphoma, diffuse large cell lymphoma and follicular lymphoma cells.

30

-106-

50. The pharmaceutical composition of claim 49, wherein  
the B cell lymphoma cells are Mucosa-Associated-Lymphoid Tissue B cell lymphoma (MALT) cells.
- 5 51. The pharmaceutical composition of claim 49, wherein  
the B cell lymphoma cells are non-Hodgkin's lymphoma  
cells.
- 10 52. A pharmaceutical composition comprising an amount of  
the oligonucleotide of claim 31 effective to prevent  
overexpression of a human IRTA protein and a  
pharmaceutically acceptable carrier.
- 15 53. A method of diagnosing B cell malignancy which  
comprises a 1q21 chromosomal rearrangement in a  
sample from a subject which comprises:
  - a) obtaining the sample from the subject;
  - b) contacting the sample of step (a) with the  
antibody of claim 43 capable of specifically  
binding with a human IRTA protein selected from  
the group consisting of human IRTA1, IRTA2,  
IRTA3, IRTA4 and IRTA5 IRTA protein on a cell  
surface of a cancer cell under conditions  
permitting binding of the antibody with human  
IRTA protein on the cell surface of the cancer  
cell, wherein the antibody is labeled with a  
detectable marker; and
  - c) detecting any binding in step (b), wherein  
detection of binding indicates a diagnosis of B  
cell malignancy in the sample.
- 20
- 25
- 30

-107-

54. The method of claim 53, wherein the B cell malignancy is selected from the group consisting of B cell lymphoma, multiple myeloma, Burkitt's lymphoma, mantle cell lymphoma, marginal zone lymphoma, diffuse large cell lymphoma and follicular lymphoma.
- 5 55. The method of claim 54, wherein the B cell lymphoma is Mucosa-Associated-Lymphoid Tissue B cell lymphoma (MALT).
- 10 56. The method of claim 54, wherein the B cell lymphoma is non-Hodgkin's lymphoma.
- 15 57. A method of treating a subject having a B cell cancer which comprises administering to the subject an amount of anti-IRTA antibody effective to bind to cancer cells expressing an IRTA protein selected from the group consisting of human IRTA1, IRTA2, IRTA3, IRTA4 and IRTA5 so as to prevent growth of the cancer cells and a pharmaceutically acceptable carrier, thereby treating the subject.
- 20 58. The method of claim 57, wherein the anti-IRTA antibody is a monoclonal antibody.
- 25 59. The method of claim 58, wherein the monoclonal antibody is a murine monoclonal antibody or a humanized monoclonal antibody.
- 30 60. The method of claim 57, wherein the anti-IRTA antibody is a polyclonal antibody.

-108-

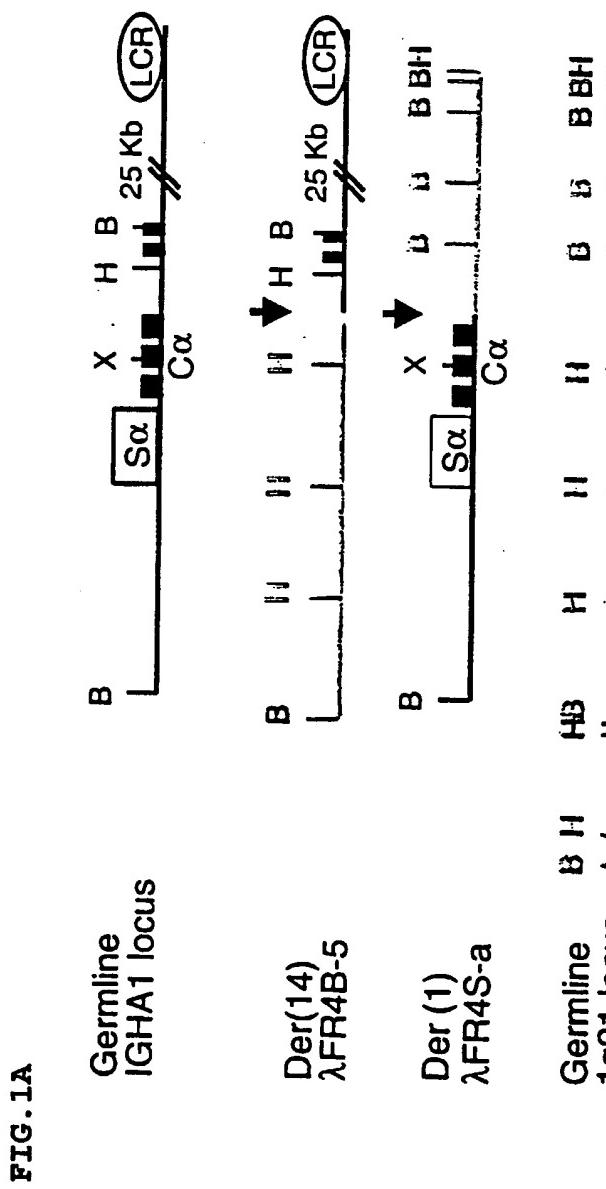
61. The method of claim 57, wherein the B cell cancer is selected from the group consisting of B cell lymphoma, multiple myeloma, mantle cell lymphoma, Burkitt's lymphoma, marginal zone lymphoma, diffuse  
5 large cell lymphoma and follicular lymphoma.
62. The method of claim 61, wherein the B cell lymphoma is Mucosa-Associated-Lymphoid Tissue B cell lymphoma (MALT).  
10
63. The method of claim 61, wherein the B cell lymphoma is non-Hodgkin's lymphoma.
64. A method of treating a subject having a B cell cancer  
15 which comprises administering to the subject an amount of the oligonucleotide of claim 31 effective to prevent overexpression of a human IRTA protein, so as to arrest cell growth or induce cell death of cancer cells expressing IRTA protein(s) and a pharmaceutically acceptable carrier, thereby treating  
20 the subject.
65. The method of claim 64, wherein the IRTA protein is selected from the group consisting of human IRTA1,  
25 IRTA2, IRTA3, IRTA4 and IRTA5 protein.
66. The method of claim 64, wherein the B cell cancer is selected from the group consisting of B cell lymphoma, mantle cell lymphoma, multiple myeloma,  
30 Burkitt's lymphoma, marginal zone lymphoma, diffuse large cell lymphoma and follicular lymphoma.

-109-

67. The method of claim 66, wherein the B cell lymphoma is Mucosa-Associated-Lymphoid Tissue B cell lymphoma (MALT).

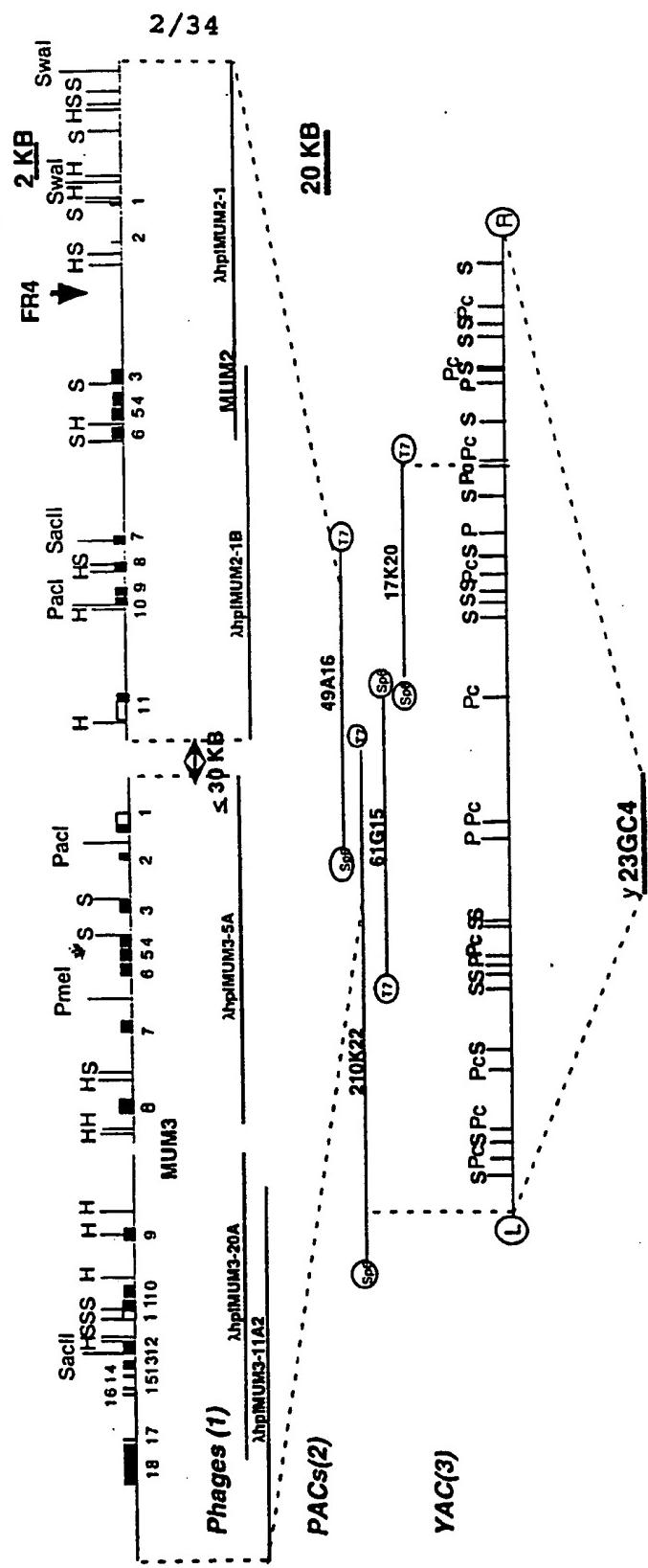
5 68. The method of claim 66, wherein the B cell lymphoma is non-Hodgkin's lymphoma.

1/34

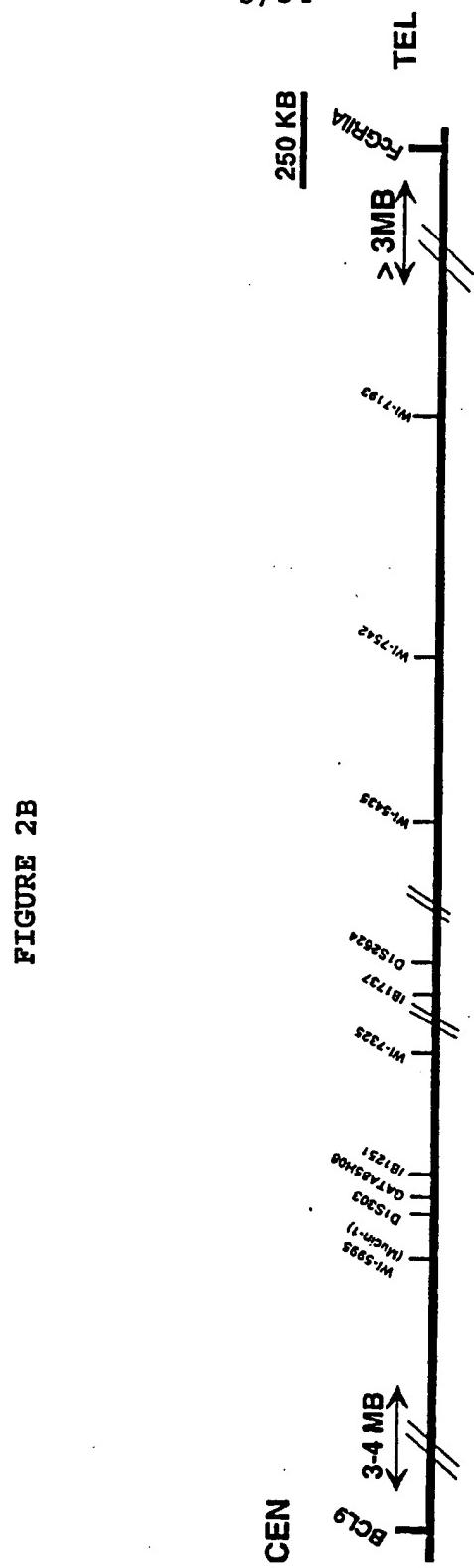
**FIGURE 1****FIG. 1B**

Chr 1. GGGCCTGACAGCAACTTCTTCTACTAGTTCATCTTAACCTTATCCCTGGTAAC"GGCGAGACAACCTGTCTTAAGTAACTGAACGGAAA  
 der14 GGCCTGACAGCAACTTCTTCTACTAGTTCATCTTAACACTGCTCTGTACCGGGCACGTGGCACAGGTGCACACTCACCTCACACCA  
 Chr 14 TCCCACTGACGCCATGAGGAAGGGCACCTCCCTTAACCAACACTGCTGTACGGGGCACGTGGCACGGTGCACACTCACACCA

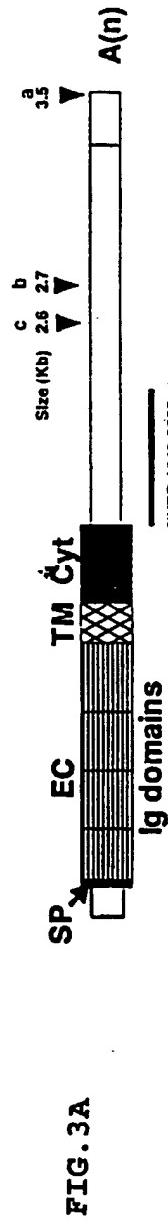
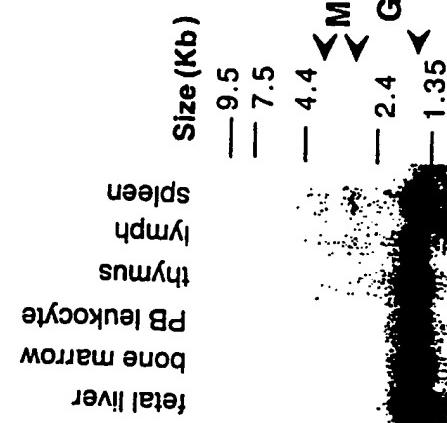
**FIGURE 2A**



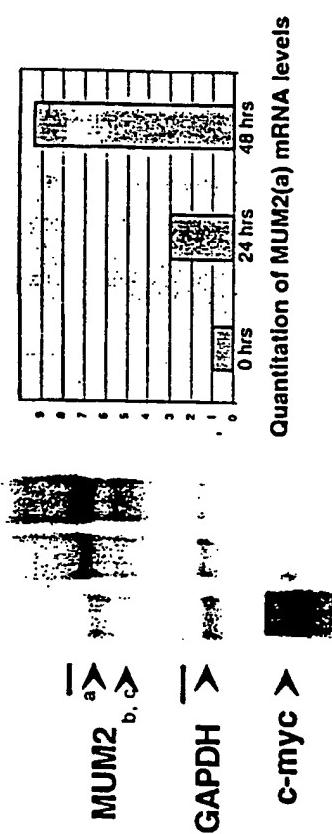
3 / 34



4/34

**FIGURE 3****FIG. 3B**

**FIG. 3C**      G0/G1 arrest →  
time(hrs)      0      24      48  
estrogen      +      -      -



Quantitation of MUM2(a) mRNA levels

FIGURE 4

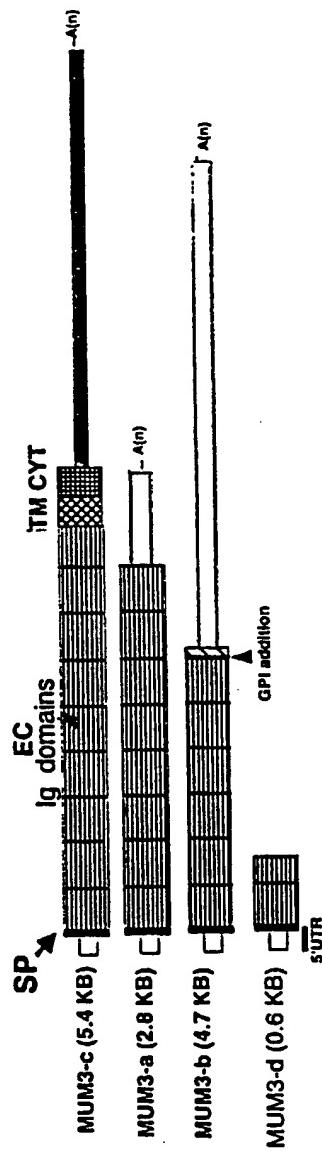
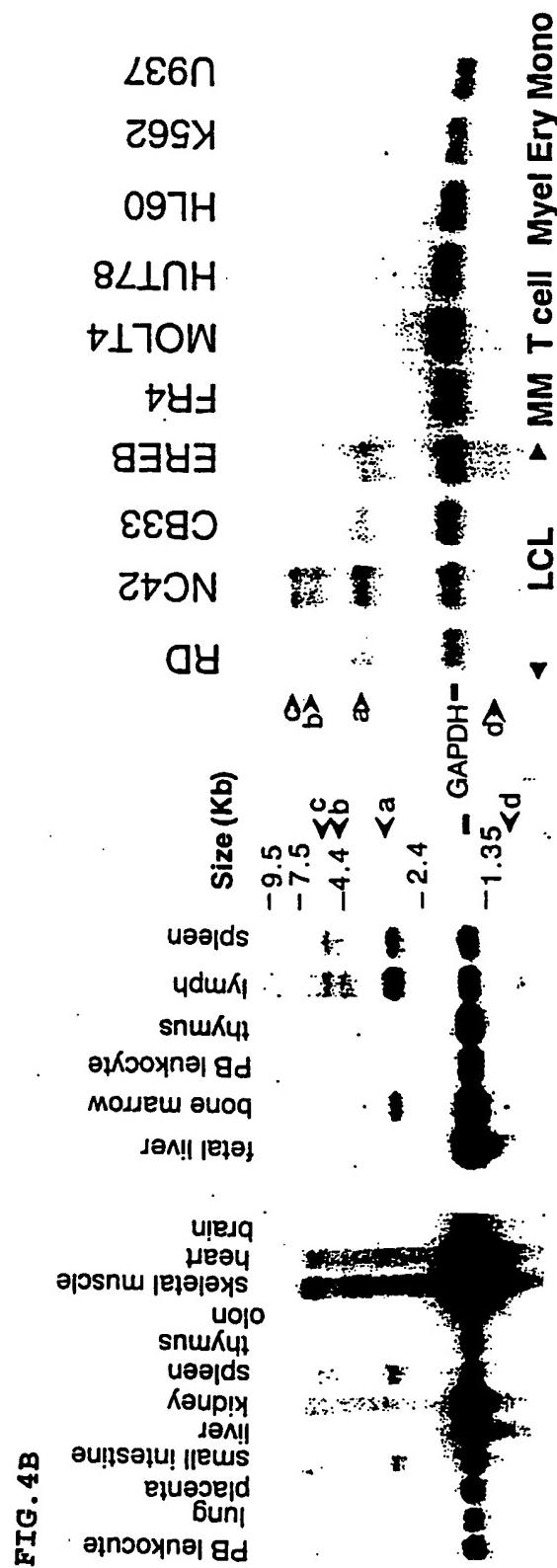


FIG. 4A



6/34

## FIGURE 5

1 CTCAACTAGCTTATGCAGAGAAGCTACTGAGCTCACTGCTGGTGTAGGCAAGTGCTGCTTGGCAA  
 M L L W A S  
 78 TCTGGCTGACCTGGCTTGTCTCCAGAACCTCTCAACCCCTGGAGCAGGCTTCCATGCTGCTGGCGTCC  
 L L A F A P V C G Q O S A A A H K P V I S V H P P W T 32  
 155 TTGCTGGCTTGTCTCCAGCTGTGGACATCTGCAAGCTGACACACAAACCTGTGATTTCGGCATCTCCATGGAC  
 T F F K G E R V T L T C N G F Q F Y A T E K T T W Y 58  
 232 CACATTCTCAAAGGAGAGAGACTGACTCTGACTTGCAATGGATTCTAGTTCTATGCAACAGAGAAAACAACATGGT  
 H R H Y W G E K L T L T P G N T L E V R E S G L Y 83  
 309 ATCATCGGCACTACTGGGAGAAAAGTTGACCCCTGACCCAGGAACACCCCTGAGGTTCGGGAACTGGACTGTAC  
 R C Q A R G S P R S N P V R L L F S S D S L I L Q A 109  
 386 AGATGCCAGGGCCGGGCTCCCCACGAAGTAACCTGTGGCTTCTCAGACTCTTAATCTGCCAGGC  
 P Y S V F E G D T L V L R C H R R R K E K L T A V K 135  
 463 ACCATATTCTGTGTTGAAGGTGACACATTGGTCTGAGATGCCACAGAAGAAGAGAAATTGACTGCTGTGA  
 Y T W N G N I L S I S N K S W D L L I P Q A S S N 160  
 540 AATATACTGGAATGAAACATTCTTCCATTCTAATAAAAAGCTGGATCTTCTATCCCACAAGCAAGTCAAT  
 N N G N Y R C I G Y G D E N D V F R S N F K I I K I 186  
 617 ACAATGGCAATTATCGATGCATTGGATATGGAGATGAGATGATGTTAGATCAAATTCAAATAATTAAAAT  
 Q E L P H P E L K A T D S Q P T E G N S V N L S C 212  
 694 TCAAGAACTATTCCACATCCAGAGCTGAAAGCTACAGACTCTCACCTACAGAGGGGAATTCTGTAACCTGAGCT  
 E T Q L P P E R S D T P L H F N F F R D G E V I L 237  
 771 GTGAAACACAGCTTCCCTCAGAGCGGTAGACACCCCACTTCACCTCAACTTCTCAGAGATGGCAGGTCACTCTG  
 S D W S T Y P E L Q L P T V W R E N S G S Y W C G A 263  
 848 TCAGACTGGAGCACGTACCCGGAACTCCAGCTCCACCGTCTGGAGAGAAAACCTCAGGATCTTGGTGTGGTC  
 E T V R G N I H K H S P S L Q I H V Q R I P V S G V 289  
 925 TGAAACAGTGGGGTAACATCCACAGCACAGTCCCTCGTACAGATCCATGTCAGCGGATCCCTGTGTCTGGGG  
 L L E T Q P S G G Q A V E G E M L V L V C S V A E 314  
 1002 TGCTCCCTGGAGACCCAGCCCTCAGGGGGCCAGGCTGTTGAAGGGAGATGCTGGCTTGTCTGCTCCGTGGCTGAA  
 G T G D T T F S W H R E D M Q E S L G R K T Q R S L 340  
 1079 GGCACAGGGATACCACATTCTCTGGCACCGAGAGACATGAGAGAGTCTGGGAGGAAAACCTCAGCGTTCCCT  
 R A E L E L P A I R Q S H A G G Y Y C T A D N S Y G 366  
 1156 GAGAGCAGAGCTGGAGCTCCCTGCCATCAGACAGGACATGAGGGGATACTGAGACAGACAACAGCTACG  
 P V Q S M V L N V T V R E T P G N R D G L V A A G 391  
 1233 GCCCTGTCCAGAGCATGGTGTCAATGTCACTGTGAGAGAGACCCAGGCAACAGAGATGGCCTTGTGCCCGGGAA  
 A T G G U L S A L I L A V A L I F H C W R R R K S G 417  
 1310 GCCACTGGAGGGCTGCTCAGTGTCTTCTCTGGCTGTCGGCCCTGCTGTTCACTGCTGGCTCGGAGGAGTCAGG  
 V G F L G D E T R L P P A P G P G E S S H S I C P A 443  
 1387 AGTTGGTTCTGGAGACGAAACCCAGGCTCCCTCCCTCCAGGCCAGGAGATCTCCATTCCATCTGCCCTG  
 Q V E L Q S L Y V D V H P K K G D L V Y S E I Q T 468  
 1464 CCCAGGTGGAGCTTCAGTGTGTATGTTGATGTACACCCAAAAGGGAGATTGGTATACTCTGAGATCAGACT  
 T Q L G E E E E A N T S R T L L E D K D V S V V Y S 494  
 1541 ACTCAGCTGGAGAAGAAGAGGAAGCTAACCTCCAGGACACTTCTAGAGGATAAGGATGTCAGTTGTACTC  
 E V K T Q H P D N S A G K I S S K D E E S \* 515  
 1618 TGAGGTAAAGACACAAACACCCAGATAACTCAGCTGGAAAGATCAGCTTAAGGTGAAGAAAGTTAAGGAGATGAA  
 1695 AGTTACGGGAACGTCCTACTCATGTAFTTCTCCCTGTCCAAGTCCAGGCCAGTCAGTCCTTGCAGGCCACCTG  
 1772 AGATGATCAACTCATCCAGTTCTAATTCTCTCATGCATATGCAATTCACTCCAGGAATACTCATCTGCTACT  
 1849 CTGATGTTGGGATGGAATGGCCTCTGAAAGACTTCACTAAATGACCAGGATCCACAGTTAAAGAGAAGACCCGTAG  
 1926 TATTTGCTGTGGGCTGACCTAATGCAATTCCCTAGGGCTGCTTTAGAGAAGGGGATAAAGAGAGAGAAGGACTGT  
 2003 TATGAAAAACAGAAGCACAAATTGCTGAATTGGGATTGCAAGAGATGAAAAGACTGGGTGACCTGGATCTCTGC  
 2080 TTAATACATCTACAACCATGGTCTCACTGGAGACTCACTTGCACTCAGTTGTTAACTGTCAGTGGCTGCACAGGCA  
 2157 CTGTGCAAAACATGAAAAGCCCCCTCACTCTGCTGCAAGCTTACACTGTCAGGATTGCAAGATTAAAGAA  
 2234 CCCATCTGGAAATGGTTACAGAGAGAGAATTAAAGAGGACATCAGAAGAGCTGGAGATGCAAGCTCTAGGCTGC  
 2311 GCTTCCAAAAGCAAATGATAATTATGTTAATGTCATTAGTGACAAAGATTGCAACATTAGAGAAAAGACACAAA  
 2388 TATAAAATTAAAACCTTAAGTACCAACTCTCCAAAACTAAATTGAACTTAAATATTAGTATAAAACTCATATAAA  
 2465 CTCTGCCTTAAATAAAAAAAAAAAAAA

7/34

FIGURE 6a

1 CCGTGCAGTGTCTGACTGTAAGATCAAGTCCAAACCTGTTGGATTGAGGAAACTTCCTTTGATCTCAGCCCCGG  
 M L L W V I L L V L A P V S G □ Q F A R T P R 22  
 61 GTGGTCCAGGTCTCATGCTGCTGGGTGATATTACTGGTCTGGCTCTGTCAGTGGACAGTTGCAAGGACACCCAG  
 P I I F L Q P P W T T V F Q G E R V T L T C K G F R F 49  
 161 GCCCATTATTTCTCCAGCTCCATGGACCACAGTCTTCCAAGGGAGAGAGAGTGACCCCTACTTGCAGGGATTTCGCT  
 Y S P Q K T K W Y H R Y L G K E I L R E T P D N I L 75  
 241 TCTACTCACCAAGAAAACAAAATGGTACCATCGGTACCTTGGAAAGAAATACTAAGAGAAACCCAGACAATATCCTT  
 E V Q E S G E Y R C Q A Q G S P L S S P V H L D F S S 102  
 321 GAGGTTCAGGAATCTGGAGAGTACAGATGCCAGGGCCAGGGCTCCCTCTCAGTAGCCCTGTGCACTGGATTTCCTC  
 A S L I L Q A P L S V E G D S V V L R C R A K A E V 129  
 401 AGCTTCGCTGATCTGCAAGCTCCACTTCTGTGTTGAGGAGACTCTGTTCTGAGGTGCGGGCAAAGGCGGAAG  
 T L N N T I Y K N D N V L A F L N K R T D F H I P H 155  
 481 TAACACTGAAATAACTATTTACAAGAATGATAATGTCCTGGCATTCTTAATAAAAGAACTGACTTCATATCCTCAT  
 A C L K D N G A Y R C T G Y K E S C C P V S S N T V K 182  
 561 GCATGTCTCAAGGACAATGGTCATATCGTGTACTGGATATAAGGAAAGTTGTTGCCCTGTTCTCCAATACAGTCAA  
 I Q V Q E P F T R P V L R A S S F Q P I S G N P V T L 209  
 641 AATCCAAGTCCAAGGCCATTACACGTCCAGTGTGAGAGGCCAGCTCCAGGCCATCAGGGAAACCCAGTGACCC  
 T C E T Q L S L E R S D V P L R F R F F R D D Q T L 235  
 721 TGACCTGTGAGACCCAGCTCTAGAGAGGTGAGATGTCCCGTCCGTTCTCAGAGATGACAGACCCCTG  
 G L W S L S P N F Q I T A M W S K D S G F Y W C K A 262  
 801 GGATTAGGCTGGAGCTCTCCCGAATTCCAGATACTGCCATGTGGAGTAAGGATTCTGAGGTTCTACTGGTGTAAAGGC  
 A T M P H S V I S D S P R S W I Q V Q I P A S H P V L 289  
 881 AGCAACAATGCCCTCACAGCGTCATATCTGACAGCCGAGATCTGGATACAGGTGAGATCCCTGCATCTCATCCTGTCC  
 T L S P E K A L N F E G T K V T L H C E T O E D S L 315  
 961 TCACTCTAGCCCTGAAAAGGCTCTGAATTGGAGGGAAACAGTGCACACTCTGAAACCCAGGAAGGATCTCTG  
 R T L Y R F Y H E G V P L R H K S V R C E R G A S I S 342  
 1041 CGCACTTGTACAGGTTTATCATGAGGGTGTCCCTGAGGACAAGTCAGTCCGCTGTGAAAGGGAGCATCCATCAG  
 F S L T T E N S G N Y Y C T A D N G L G A K P S K A V 369  
 1121 CTTCTCACTGACTCACAGAGAACTACTACTGACACAGTCACAACTGGCTTGGCCAAGCCCAGTAAGGCTG  
 S L S V T P V S H P V L N L S S P E D L I F E G A 395  
 1201 TGAGCCTCTAGTCACTGTTCCGTTCTCATCTCTCAACCTCAGCTCTCTGAGGACCTGATTTGAGGGAGCC  
 K V T L H C E A Q R G S L P I L Y Q F H H E D A L E 422  
 1281 AAGGTGACACTTCACTGTGAAGGCCAGAGGGTCACTCCCCATCTGTACCAAGTTCATCATGAGGATGTCCTCTGGA  
 R R S A N S A G G V A I S F S L T A E H S G N Y Y C 449  
 1361 CGGTAGCTGGCCAACCTCTGCAAGGGAGTGGCATCTGACTGCAAGGACTTCAGGGAACTACTGCA  
 A D N G F G P Q R S K A V S L S I T V P V S H P V L 475  
 1441 CAGCTGACAATGGCTTGGCCCCAGCCAGTAAGGGCGTGAGGCTCTCCATCACTGTCCTGTCTCATCTGTCTC  
 T L S S A E A L T F E G A T V T L H C E V Q R G S P Q 502  
 1521 ACCCTCACTCTGAGGCCCTGACTTTGAAGGAGCCACTGTGACACTTCACTGTGAAGTCCAGAGGGTCCCCACA  
 I L Y Q F Y H E D M P L W S S S T P S V G R V S F S 529  
 1601 AATCCCTAACCGTTTATCATGAGGACATGCCCTGAGGAGCTCAACACCCCTGTGGGAAGACTGTCCTTCAGCT  
 S L T E G H S G N Y Y C T A D N G F G P Q R S E V V 555  
 1681 TCTCTCTGACTGAAGGACATTCAAGGGAAATTACTACTGCAAGCTGACAAATGGCTTGGCTCCAGGCCAGTGAAGTGGT  
 S L F V T V P V S R P I L T L R V P R A Q A V V G D L 582  
 1761 AGCCTTTTGTCACTGTTCCAGTCTCGGCCCCATCTCAGGGCTTCCAGGGCTGGAGGGGGACCT  
 L E L H C E A P R G S P P I L Y W F Y H E D V T L G S 609  
 1841 GCTGGAGCTCACTGTGAGGCCAGAGGGCTCTCCCCAATCTGTACTGGTTTATCATGAGGATGTCACCCCTGGGA  
 S S A P S G G E A S F N L S L T A E H S G N Y S C E 635  
 1921 GCAGCTAGCCCCCTCTGGAGGAGAAGCTTCTCAACCTCTCTGACTGCAAGAACATTCTGAAACTACTCATGTGAG  
 A N N G L V A Q H S D T I S L S V I V P V S R P I L T 662  
 2001 GCCAACATGGCTAGTGGGCCAGCACAGTCACAAATATCACTCACTGTTATAGTCCAGTATCTGCTCCATCCCTCAC  
 F R A P R A Q A V V G D L L E L H C E A L R G S S P 1689  
 2081 CTTCAAGGGCTCCAGGCCAGGTGTGGGGACCTGCTGGAGCTCACTGTGAGGCCCTGAGAGGCTCTCCCCAA  
 L Y W F Y H E D V T L G K I S A P S . G G G A S F N L 715  
 2161 TCCCTGACTGGTTTATCATGAAGGATGTCACTGGTAAGATCTCAGCCCCCTCTGGAGGGGGCCCTCTCAACCTC  
 S L T T E H S G I Y S C E A D N G L E A Q R S E M V T 742  
 2241 TCTCTGACTACAGAACATTCTGAAATCTACTCTGTGAGGAGACAATGGTCTGTGAGGCCAGCGAGTGAAGATGGTAC  
 L K V A G E W A L P T S S T S E N \* 759  
 2321 ACTGAAAGTTGCAGGTGAGTGGGCCCTGCCACCAAGCACATCTGAGAAGTCACTGTCCTGTTCTCCCTGCAGCTGA  
 2401 AAATGGAGCCACAGAGCTCTCAGGGCTGTTGCTTGAGGATCTCAGCCCCCTCTGGAGGGGGCCCTCTCAACCTC  
 2481 AAAGTCTGGATCTCTGTGGTATGGTCCAGGAATCTGATGTTCTCCAGCTTCTGAGATGATCAAAGCACCTC  
 2561 ACTAAAAATGCAAAATAAGACTTTAGAACATAACTATATTCTGACTGAAATTATTACATGAAAATGAAACCAAGA  
 2641 ATTCTGAGCATATGTTCTCTGCCGTAGAAGGATTAAGCTTGTGAGGGCCAGCGAGTGAAGATGGTAC  
 2721 GCCTCTACTCTGAGTCTTTCACTACTGGGATGTAATGTTCCACATTAAATCCTATGACTTCAAGAA  
 AAAAA

8/34

## FIGURE 6b

1 CGGTGCACTGTCCTGACTGTAAGATCAAGTCAAACCTGTTGGATTGAGGAACCTGTTTGATCTCAGCCCTTG  
   M L L W V I L E L V L A P V S G Q F A R T P R 22  
 81 GTGGTCAGGTCTTCATGCTGCTGGGGTCAATTACTGGCTCTGCCCTGTCAGTGGACAGTTGCAAGGACACCCAG  
   P I I F L Q P P W T T V F Q G E R V T L L T C K G F R F 49  
 161 GCCCATTTTCTCCAGCCTCCATGGACCACAGTCCTCAAGGAGAGAGTGCACCTCACTTGCAGGGATTTCGCT  
   Y S P Q K T K W Y H R Y L G K E I L R E T P D N I L 75  
 241 TCTACTCACACAGAAACAAATGGTACCATCGGCTACCTGGAGAAGAAATACTAGAGAAAACCCAGACAATATCCTT  
   E V Q E S G E Y R C Q A O G S P L S S P V H L D F S S 102  
 321 GAGGTTCAAGGAATCTGGAGAGTACAGATGCCAGGCCAGGGCTCCCTCTAGTAGGCCCTGTGCACTTGGATTTC  
   A S L L Q P M P L S V F E G D S V U L R C R A K A E V 129  
 401 AGCTTCGCTGATCTGCAAGCTCCACTTCTGTTGAAGGAGACTCTGTTCTGAGGTCGGGGCAAAAGGGGGAG  
   T L N N T I Y K N D N V L A F L N K R T D F H I P H 155  
 481 TAACACTGAATAATCACTTACAAGATCATACTGTCCTGCAATTAAAGAACACTGACTTCCATATTCTCAT  
   A C L K D N G A Y R C T G Y K E S C C P V S S N T V K 182  
 561 GCATGTCAGGACATGGTCATATCGCTGACTGGATAAAAGGAATTGTTGGCCCTGTTCTCCAAATACAGTC  
   I Q V E P F T R P V L R A S S F Q P I S G N P V T L 209  
 641 AATCCAAGTCCAAGGCAATTACAGCTCAGTGCAGGCCAGCTCCAGGCCATCAGCAGGGAAACCCAGTGACCC  
   T C E T Q L S L E R S D V P L R F R D D Q T L 235  
 721 TGACCTGAGGACCCAGCTCTCTAGAGGGTCAGATGCCCCCTCCGTCAGAGATGACCCAGACCTCG  
   G L G W S L S P N F Q I T A M W S K D S G F Y W C K A 262  
 801 GGATTAGGCTGGAGCTCTCCCCGAAATTCCACATTACTGCCATGTTGGACTAAAGATTGAGGGTTCTACTGGTAAAGGC  
   A T M P H S V I S D S P R S W I Q V Q I P A S H P V L 289  
 881 AGCAACAAATGCTCAGCGTCATATCTGACAGCCAGAGTCCTGACAGTCCAGATCCCTGCATCTCATCTGTCC  
   T L S P E K A L N F E G T K V T L H C E T Q E D S L 315  
 961 TCACTCTAGCCCTGAAAAGGCTCTGAATTGAGGAAACCAAGGGACACTCTACTGTGAACCCAGGAAGATCTCTG  
   R T L Y P H E G V P L R H K S V R C E R G A S I S 342  
 1041 CGCACTTGTACAGGTTTATCATGAGGTGTCCTGGACTAACAGTCAGTCTCCGTGAAAGGGGAGCATCCATCAG  
   F S L T T E N S G N Y Y C T A D N G L G A K P S K A V 369  
 1121 CTCTCTGACTACAGAGAAATCAGGAAACTACTACTGCACTGACAATGGCCTGGGCCAACCCCAGTAAGGTG  
   S L S V T V P V S H P V L N L S S P E D L I F E G A 395  
 1201 TGACCCCTCTAGTCAGTGTCCCCCTCTCATCTGTCAGCTCTCCGTCAGGACCTGATTTTGAGGGAGCC  
   K V T L H C E A O R G S S P I L Y Q F H E D A A L E 422  
 1281 AAGGTGACACTTCAGTGAAGGCCAGAGAGGTTCACTCCCATCTGACCATGTTCTCATCATGAGGATGCTGCCCTGGA  
   R R S A N S A G G V A I S F S L T A E H S G N Y Y C T 449  
 1361 GCGTAGGTGGCCAATCTGCAAGGGAGGTGGCCATCAGCTCTCTGACTCAGACAGCAATTAGGGAAACTACTGCA  
   A D N G F G P R S K A V S L S I T V P V S H P V L 475  
 1441 CAGCTGACAATGGCTTGGCCCCAGCCAGTAAGGGCTGAGCCTCTCCATCACTGTCCTGTCATCTGTCCTC  
   T L S S A E A L T F E G A T V T L H C E V Q R G S P Q 502  
 1521 ACCCTCAGCTGTCAGGCCCTGACTTTGAAGGAGCCACTGACACTTCAGTGTGAGTCAGAGAGGGTTCCCCACA  
   I L Y Q F Y H E D M P L X S S S T P S V G R V S F S 529  
 1601 AATCTTACACAGTTTATCATGAGGACATGCCCCCTGGAGCAGCTAACACCCCTCTGGAAGAGTGTCTCTCAGT  
   S L T E G H S G N Y Y C T A D N G F G P Q R S E V V 555  
 1681 TCTCTGACTGAAGGACATTCAAGGAAATTACTACTGCAAGCTGACAGCAATTGGCTTCCCAGGGCAGTGAAGTGGC  
   S L F V T G R C W V L A S H P P L A E F S L T H S F K 582  
 1761 AGCCCTTTGTCACTGGTAAGTGTGGTTCTGCCAGTCACCCACCCCTGGCTGAGTTCTCTCACCCATTCTTAA  
   N L F A L S S C T C - seop 592  
 1841 AAATCTGTTGACTCTCCCTAATCACTTAATCTCTCTCTCAACTAACTACTAGCTGG  
 1921 GTTTCGTAATCTCATGCTGGCTCAGCCAGACCTTAAACAGCTCAGTAGATCTCCCTTACCAAATGAAATT  
 2001 TATTATGTGTTCTCTCATCTGTATGTTCTCAAGTAGCAGGCCAATTCTCTGATGCAAGGGCTGTCCTACT  
 2081 TCTCTGACTGACATTCTCATATTAACCTAGCTACAAAGCACAGTCTTATAGATAAATATGGTCAAGACCTTAAATTCTCCA  
 2161 AAGGATTCCAATCTTATGGTAGATTGGAGAAAGCTGCTGGTGAACAAAGGGGAAATGGCTCCCTAGGAACCAACTCC  
 2241 TCAACTCTGGAGTTTATGATCTGCTCTGTTCTCTGCTAACTGCTAAATCACTGATCTATCTTATGTTATTTAA  
 2321 ACTATTGTTGAGATTGACATCATCTGAGAAACCTGTCAGAAATTGAGCTTCTGAGCTTCTTCTGAGGCTAA  
 2401 GTTGTCTCTGTTTACTGAAATCTTGTATTAGAAACTGGGGAAAGTTTACTCTTCTGAGTAAAGGGTTTTAAAGGT  
 2481 GTAGAGAAAAATCTGAGGCTGAGCTGAGACATGCCCTAGCATAACTGTTGAGTAAAGGGTTTTAAAGGT  
 2561 GATGTTCTGAGACTACTCCAAAGTCAGGCCAACTTACTGAGGAACTTCTAGACTTCATCTGCACTCCCATAC  
 2641 TCTCTTATCTCATCTGTTACTCTCTCATATGAGGACATCTGAGGACATCTTCTGTTCTGACTGTC  
 2721 CCCTAATGCCAGTAGAATGTAAGCTTATGAGAACAGAACACTGCATCCATCTGCTTCTCACACATCTGCTACT  
 2801 CAGTGTGGCACACAGTAGGCTCTGACTAACCTGTTATTAGTGTGAGCAGATGATGACAAGATGATAAGAGGGGA  
 2881 TTTAAAGGAAATCTGAGGCAAGGCCAGAGGAAAAAAACAAAGCTTATTAGAATGAAATACCAATTGAGGAGCTA  
 2961 AGATAGATTGGATATCTTGGAAACCATTTATTGATGAGAACCTTCTGAGGAAACAAATACAGAAATGCAAGTAGAA  
 3041 AGATACAGAAATTAAGGCAAAGTAAATGAAATGAAATGAGAACATGAGGATTTGTCTGATCCAGTTATGTGATAA  
 3121 ATGGAGACCCCTCAGAAATTGAGGAAAGCTGAGGAAATGAAACTCTGAGGAAATTGAGTGAATGTTGGGAAAGTAAAGAA  
 3201 ACTGGAATATGAGATGCAAAATATGTTGAGCCTTGTGACTTGTGAGGTTAAATATATATGTTGCTCTGAT  
 3281 ATGGGGAAAAAAGGAGTGTCTCAGAAAGAAAAACATCAAGTTAGTGTGACTTGTGAGTCACTGAGTACCAAGAG  
 3361 AGAGGGCCAGACTGGACCTGGAGGGGAAGAATACCCAAATTCTTATCTGAACTTCACTTCAAGGACATTGTCAAAAA  
 3441 TACAGGAGATTCAAGGAAACTTGAGAAATGCAACTGAGCTTCTGAGGAAACACCTAATGCAAAATCTAGGCCAACAAAGATGT  
 3521 AAATGAATATAAGGACTCATATGAGGAAACCCGCTTGTACTGCTCTCAACCTGGCCGATATTAGACTCTGCAAA  
 3601 GACCTTGTAAAAGCTCACACATGACTCTGCAAGGCCCTCTCCAGACTAACTCAATTCAAGAACCTCACAGATGGGGCC  
 3681 ACAGAAATCTGAGTATTGTTGACACAACTCTCAAGTGAGAAATTGAGCTGAGGAAACACTGATTAGATATAGA  
 3761 AACAAAGGCTAATCACTGTGAGAATTATGGTCAGCAAGAAAGTAACTATTGAACTGAAATGTTAAAGGAA  
 3841 GTAAACAAAGGAAATTAGTGTAGGAGGAGGAGGAAGTAAAGGAACAACTCATTTCTCATGATTATTATTTAGAGTA  
 3921 ATTGTGAGTTATTCTACAATTCAAAGGAAATGGACTTTAAAGGAAATGGACTTTAAAGGAAATGGCTTCTGCT  
 4001 AATGTTCTGAAATCTGTCAACAGTACTCATCTTAAATGGCTTATACCTCTACTAAAATCTGAGGAAACCAACTA  
 4081 GTAGCCCTGAGAGTCACATGGAGGAGCAACTGCTTGTGAGCTTCTGGCTGGCAAGGAGTGTGAGGAGCTGACAAAAAA  
 4161 AATAATAAAATAAAACCTGTGCTTGTGATATGATCACAAATGATCACGGGAAACAGGAAACAGAAACTCTCATACGCCATT  
 4241 TTACAACTGTAAATTGGTCAACCTTTCGCTCTTAACTGACACATTGTAATTGTATATTTATGGAGGACAGTTGAT  
 4321 ATTCTGATATACATGTTATGAACTGAGGAAATGGATTTAGGATTTAGTGTACCCATCTCATGCTTCTGAT  
 4401 TTGGAAATAAATCTCAAACCTCAAACCCAAAAAAGGAAATGGACTTTAAAGGAAATGGCTTCTGCT

9/34

## FIGURE 6C-1

1 CGGTGCAAGTGTCTGACTGTAAAGATCAAGTCCAAACCTGTTGGATTGAGGAACCTTTCTTGATCTCAGCCCC 22  
     M L L W V I E L V L A P V S G Q F A R T P R  
 21 GTGGTCCAGGTCTTCATGCTGCTGGGTGATAATTACTGCTCTGGCTCTGTCAGTGGACAGTTGCAAGGACACCCAG 22  
     P I I F L Q P P W T F Q G E R V T L T C K G F R F  
 161 GCCCATATTTCCTCCAGCCTCCATGGACCACAGTCTCCAAGGGAGAGAGAGTGAACCTCAGTCAAGGGATTGCT 49  
     Y S P Q K T K W Y H R Y L G K E I L R E T P D N I L  
 241 TCTACTCACCACAGAAAACAAAATGGTACCATCGTACCTGGAAAGAATACTAAGAGAAAACCCAGACAATATCCTT 75  
     E V Q E S G E Y R C Q A Q G S P L S S P V H L D F S S  
 321 GAGGTTCAAGGAACTCGAGAGTACAGATGCCAGGGCTCCCTCTCAGTAGCCCTGTGCACTTGGATTTCCTC 102  
     A S L I L Q A P L S V F E G D S V V L R C R A K A E V  
 401 AGCTTCGCTGATCTGCAAGCTCCTGTTGAAAGGAGACTCTGGGTTCTGAGGTGCCGGCAAAGGGGAAG 129  
     T L N N T I Y K N D N V L R F L N K R T D F H I P H  
 481 TAACACTGAATAATACTATTACAAGAACATGATAATGTCCTGGCATTCCTTAATAAAGAACATGACTTCCATATTCTCAT 155  
     A C L K D N G A Y R C T G Y K E S C C P V S S N T V K  
 561 GCATGCTCAAGGACAATGGTGCATTCGCTGACTGGATAAGGAAGTTGCTGCCCTGTTCTCCAAATACAGTCAA 182  
     I O V Q E P F T R P V L R A S S F Q P I S G N P V T L  
 641 ATCCAAGTCCAAGAGCCATTACAGTCCTGAGAGCCAGCTCCAGGCCCACATCAGGGGAACCCAGTGACCC 209  
     T C E T Q L S L E R S D V P L R F R F D Q T L  
 721 TGACCTGTGAGACCCAGCTCTCTAGAGAGGTCAGATGTCCTCCGCTCCGCTCTCAGAGATGACAGACCCCTG 235  
     G L G W S L S P N F O I T A M W S K D S G F Y W C K A  
 801 GGATTAGGCTGGAGTCTCTCCCGAATTTCAGATTACTGCCATGTGGAGTAAGATTCAAGGGTCTACTGGTGAAGGC 262  
     A T M P H S V I S D S P R S W I Q V Q I P A S H P V L  
 881 AGCAACAATGCTCACAGCGTCATATCTGACAGGGAGATCTGGATACAGGTGCAGATCCCTGCATCTCATCTGTCC 289  
     T L S P E K A L N F E G T K V T L H C E T Q E D S L  
 961 TCACTCTCAGGCTGAAAAGGCTGAAATTGAGGAACCAAGGTGACACTTCACTGTGAAACCCAGGAAGATTCTG 315  
     R T L R F Y H E G V P L R H K S V R C E R G A S I S  
 1041 CGCACTTGTACAGGTTTATCATGAGGTGTCCTCTGAGGCAAGACTGACTCCGCTGTAAGGGGAGCATCCATCAG 342  
     F S L T T E N S G N Y C T A D N G L G A K P S K A V  
 1121 CTCTCACTGACTACAGAGAACATTCAAGGAACACTACTGCAAGCTGACAATGCCCTGGCCCAAGCCAGTAAGGCTG 369  
     S L S V T V P V S H P V L H L S S P E D L I F E G A  
 1201 TGAGCCTCTCACTGTCTCCCTGTCATCTCTGCAACACTCAGCTCTGAGGACCTTGTGAGGATGCCCTGGGCC 395  
     K V T L H C E A Q R G S L P I L Y Q F H E D A A L E  
 1281 AAGGTGACACTTCACTGTGAAGGCCAGAGAGGTCACTCCCATCTGTACCAAGTTCATGAGGATGCCCTGGGA 422  
     R R S A N S A G G V A I S F S L T A E H S G N Y C T  
 1361 GCGTAGGTGCCAACCTCTGAGGGAGTGGCCACTCAGCTCTCTGACTGCAAGAGCACTAGGGAAACTACTGCA 449  
     A D N G F G P Q R S K A V S L S I T V P V S H P V L  
 1441 CAGCTGACAATGGCTTGGCCCTCACGGCAGTAAGGGGTGAGGCCCTCTCCATCACTGTCCCCTGTCATCTGTCTC 475  
     T L S S A E L T F E G A T V T L H C E P R Q G S P Q  
 1521 ACCCTCAGCTCTGCTGAGGCCCTGACTTTGAAGGGGCCACTGTGACACTCAGCTGAACTCAGAGAGGTTCCCAACA 502  
     I L Y Q F Y H E D M P L W S S S T P S V G R V S F S F  
 1601 AATCCTTACCAAGTTTATCATGAGGACATGCCCTGTGGAGCACCTGACTCAACACCCCTGTGGGAAGAGTGTCTCAGCT 529  
     S L T E G H S G N Y C T A D N G F G P Q R S E V V  
 1681 TCTCTGACTGAAGGACATTAGGGAAATTACTGCAAGCTGACAATGCTTGTGCTCCAGGGCAGTGAGTGGTG 555  
     S L F V T V P V S R P I L T L R V P R A Q A V V G D L  
 1761 AGCCTTTTGTCACTGTCTGCTGCCCCATCCCTCACCCCTAGGGTTCCAGGGCCAGGGCTGTGGGGGGACCT 582  
     L E L H C E A P R G S P P I L Y Q F H E D V T L G S  
 1841 GCTGGAGCTTCACTGTGAGGCCCTGAGAGGCTCTCCCAATCTGTACTGGTTTATCATGAGGATGTCACTCTGGGA 609  
     S S A P S G G E A S F H S L T A E H S G H Y S C E  
 1921 CGAGCTGAGCCCCCTCTGGAGGAGAACCTTCACTCTGACTGCAAGAACATTGAAACTACTCATCTGAG 635  
     A N N G L V A Q H S D T I L S V I V P V S R P I L T  
 2001 GCCAACATGGCTAGTGGCCCAAGCACAGTGACAAATATCACTCACTGTTATAGTCCAGTATCTGTCCCACCTC 662  
     F R A P A R A Q A V V G D L E L H C E A L R G S S P I  
 2081 CTCAGGGCTCCAGGGCCCAAGGCCCTGGTGGGGACCTCTGGAGGCTCACTGTGAGGCCCTGAGAGGCTCTCCCCA 689  
     L Y W F H E D V T L G K I S A P S G G G A S F N L  
 2151 TCCTGTACTGGTTTATCATGAGGATGTCAACCTGGTAAGATCTGAGCCCCCTCTGGAGGAGGGCTCTTCAACCTC 715  
     S L T T E H S G I Y S C E A D N G L E A Q R S E M V T  
 2241 TCTCTGACTACAGAACATTCTGGAAATCTACTCTGTGAGGCAAGAACATTGGTCTGGAGGCCAGCAGTGAGATGGTGAC 742  
     L K V A V P V S R P V L T L R A P G T H A A V G D L E  
 2321 ACTGAAAGTTGCAAGTTCGGTCTGGCCCTGCTCACCCCTCAGGGCTCCGGGACCCATGCTGGTGGGGGACCTGC 769  
     E L H C E A L R G S P L I Y R F F H E D V T L G N  
 2401 TGGAGCTTCACTGTGAGGCCCTGAGAGGCTCTCCCTGACCTCTGAGGGTTTCTCATGAGGATGTCACTCTGGCT 795  
     R S S P S G G A S L N L S T A E H S G N Y C E A D  
 2481 AGGTGCTCCCTCTGGAGGAGCGCTCTAAACTCTCTGACTGCAAGAGCACTGGAAACTACTCTCTGAGGCCGA 822  
     N G L G A Q R S E T V T L Y I T G L T A N R S G P F A  
 2561 CAATGGCTGGGGCCCAAGCGCAGTGAGAACATGACTTATACACAGGGCTGAGGCCAACAGAGTGGCCCTTTG 849  
     T G V A G G I L S I A G L A A G A L I I Y C W I S R  
 2641 CCACAGGAGTCCGGGGCCCTGCTCAGCACAGGCCCTGCTGGGGGGACTGCTGCTCTACTGCTGGCTCCGAGA 875  
     K A G R K P A S D P A R S P S D S D S Q E P T Y H N V  
 2721 AAAGCAGGGAGAAAGCTGCAACACTGTGACACTAATGCAATTCAGACTGGCAACTCCAAAGGCCACCTACACATGT 902  
     P A W E E L Q P V Y T N A N P R G E N V V Y S E V R I  
 2801 ACCAGCCTGGGAAGAGCTGCAACACTGTGACACTAATGCAATTCAGAGGAGAAAATGTGGTTACTCAGAAGTACGGA 929  
     I Q E K K H A V A S D P R H L R N K G S P I I X S  
 2881 TCATCCAAGAGAAAAAGAACATGCACTGGCCCTGACCCAGGCATCTCAGGAACAAGGGTCCCCCTATCATCTACTC 955

10/34

## FIGURE 6c-2

E V K V A S T P V S G S L F L A S S A P H R • stop 977

```

2961 GAAGTTAAGGTGGCGTCACCCGGTTCCGATCCCTTCTGGCTCAGCTCACAGATGAGTCCACACGTC
3041 TCTCCAATCTGCTGTTCAAGCCTCTGCACCCAAAGTCCCCCTGGGGAGAACGAGCATGAAAGTGGAAAGATTAGGCT
3121 GCCCCAACCATATCTACTGGCTTACATGCTCATTCTCAGTCTGACAGAACATGCAGGGCCCTGCTGGACTG
3201 TCACCTTTCCAGTTAAAGCCCTGACTGGCAGTTTAAATCCAGTGGCAAGGTGCTCCACCTCCAGGGCCACAGCAC
3281 ATCTCTGGATTCCTAGTGGCTTCAGCTGGTCTGACTGCTCATCACACCCCCACAGAGGGGTC
3361 TTACACACAAAAGGGAGGTGGGCTTCAGGAGATGCCGGCTGGCTAACAGCTCAGGTGCTCTAAACTCCGACACAG
3441 AGTTCTGCTTGGGATGCTTCAATTGCTCATCAGCTGCTGGGCTACTGCAGTGTGCTGCCAATGGGACAG
3521 CACACACCTGTCACATGGGACATGATGGTCTCCACGGGCTGCATTCAACTCCACCTGCTCAAAC
3601 CTAAGGTGGCACTTGACACCAAGGAACATCTCTCTGCTCATGTGTCAGTGTCTACCTGCCAAATAAGTGGCTTCA
3681 TACACCAAGTCCGAATTCTCCCATCTAACAGAAGTAACCCACCAAGTCAGGCCAGGAGGACAGGGGTGCAGACA
3761 GAACACATACTGGAACACAGGAGGTGCTCAATTACTTTGACTGACTGACTGAATGAATGAATGAGGAAGAAAAC
3841 TGTGGTAATCAAACCTGGCATAAAATCAGTCACCTGGCTTAGGAAATGGGAGGTATTCTGGCTTCTAAGAAAACAAACG
3921 GAAGAGAAGGAGCTGGATGAAAGAACCTGTCAGCAAGAACAGGGCTCTCACACTTTATGTGCTTGTGGATCACCT
4001 GAGGATCTGAAAATACAGATACTGATTCAAGTGGCTGTGAGACCTGAGACTGCCATTCAACATGTCCAGGGG
4081 ATGCTGATGCTGCTGCCCTGGGACTGCTGACATGCTGAAGGGCTTATAGGCTTCAGCAGGGCCATGGAGAGGGA
4161 ATGTTGGCTCTGGCTGCCCTGGGACTGCTGACATGCTGAAGGGCTTACAGGGCCATGGGAGGCTTGGGCCACAGCAC
4241 CACCAAGCTGCTGGTGTGAGAGGCTCTCTGTGACATGTTGGCTTACATCAGGCCACCCCTGGGAAGGGAAAGTAGC
4321 TGCCACTATCTTGTCTCCCATCTGCAAGGCTCACACTTCCCAGTAAAAGGGTGAATGTATATAACCTGAGCCCTCTCC
4401 ATTCAAGAGTTGGTCTCCCATCTGAGCAATGGGATGTTCTGTTCCGCTTTATGATATCATCAGTCTTATCTGATC
4481 TTTGCTCCAGTGGATGTAAGTGTGACTTTAACGCCCCACGGGCTGAAATAATCTTCAAGGGCATGGAAAGC
4561 TCACTCCACCTGAACCAGGTTTCACTGCTTCAAGTGTCAAGGGCTTGGCCAGATAGACAGGGCTGACTCTGCTGCC
4641 CAACCTTCAAGGAGGAACCCAGACACCTGAGACAGGAGCTGTATGCAGCCCTGCAAGGACAAGGCTG
4721 GAGGCATTGTCATCACTACAGATATGCAACTAAAATAGACGTTGGAGCAAGGAAATGCAATTCCACCGAGGCCCTTT
4801 TTAGGCTTAGTTGAAGTCAGAAGGACAGCAGCAAGCATGGCTCAGGATTAAGGAAAAAAATCTGTCACAGTCTGTT
4881 CTGGAGGTACATCACAACAAAGCTCACGGCCATGCAAGTCTGAGAAGGTGGAGGCACCGGCTCAAAAGGAAAATT
4961 TAGAATTCTCATGGGAGAGTAAGGTACCCCATCCAGAACATGATAACTGCACAGTGGAGAACAAACTCCACCCATAAT
5041 GTGGGTGGACCCATCCAGTGTGAGGGCTGAATGTAACAAAAGGGCTTATCTTCCCTCAAGTAAGGGGAACCTCT
5121 GCTTTGGCTGGACATAAGTTCTGCTTCAAGCAGCAACTGAAAAATGGCTTCTGGGTCTTGAGCTTGTGGC
5201 ATATGGACTGAAAGAAACTATGCTATGGATCTCTGGATCTCCAGCTGCTGACTGAGATCTTGAGATATGTCAAGCCT
5281 CTACACTCACAAGAGCTAATTCAATTAACCAATCTTC

```

11/34

FIGURE 7

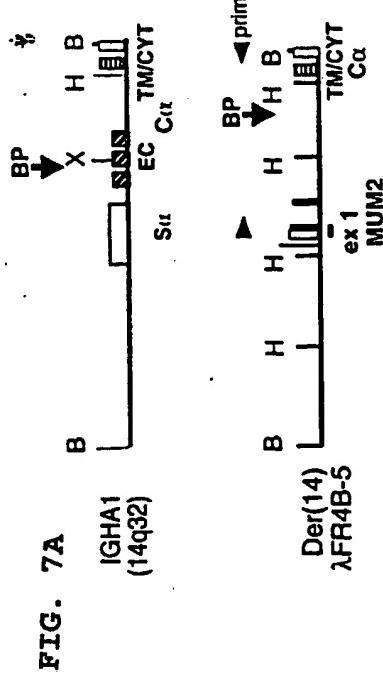


FIG. 7B

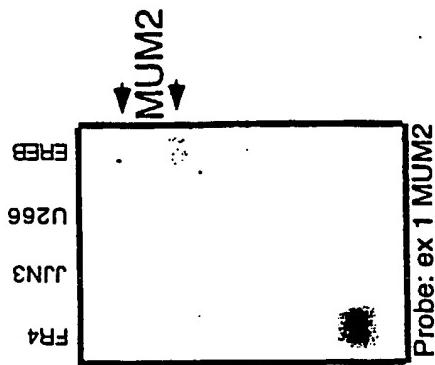
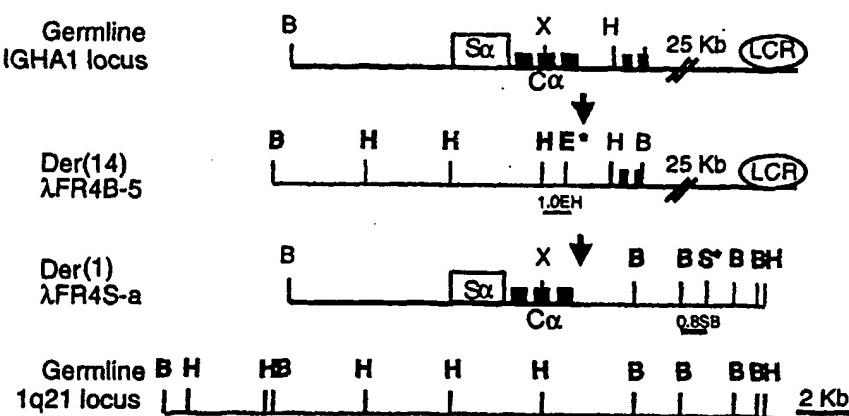


FIG. 7C



12/34

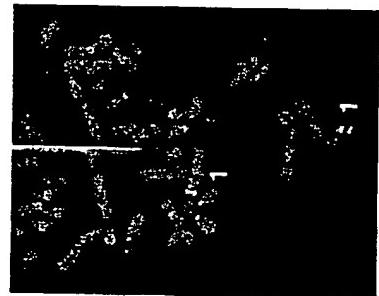
FIGURE 8A



13 / 34

FIGURE 8  
FIG. 8B

Chr	14	TCCCACTGACAGCAACTTTCTCTACTAGTCATCATTAACTTGTAACTGGTAACCTGTTAAGTAACCTGAAGGGAA
Der	(14)	GGCCTGACAGCAACTTTCTCTACTAGTCATCATTAACTTGTAACTGGTAACCTGTTAAGTAACCTGAAGGGAA
Chr	1	GGCCTGACAGCAACTTTCTCTACTAGTCATCATTAACTTGTAACTGGTAACCTGTTAAGTAACCTGAAGGGAA
Chr	1	TCCC ACTGAGCA --- GGAAGG ATCTTGTTATCTGGTAACCTGGGAGACAACCTGTTAAGTAACCTGAAGGGAA
Der	(1)	TCCC ACTGAGCA --- GGAAGG ATCTTGTTATCTGGTAACCTGGGAGACAACCTGTTAAGTAACCTGAAGGGAA
Chr	14	TCCC ACTGAGCA --- GGAAGG ATCTTGTTATCTGGGACCTCTCCCTTAACCACACTGCTCTGTAACTGGGACAGGTGCACACTCAC



8C

14/34

FIGURE 9

FIG. 9A

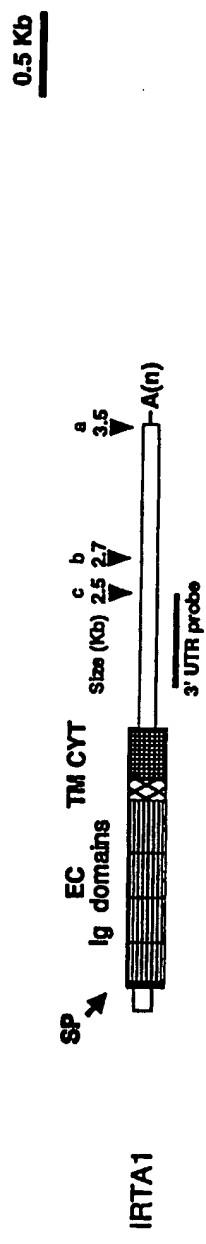
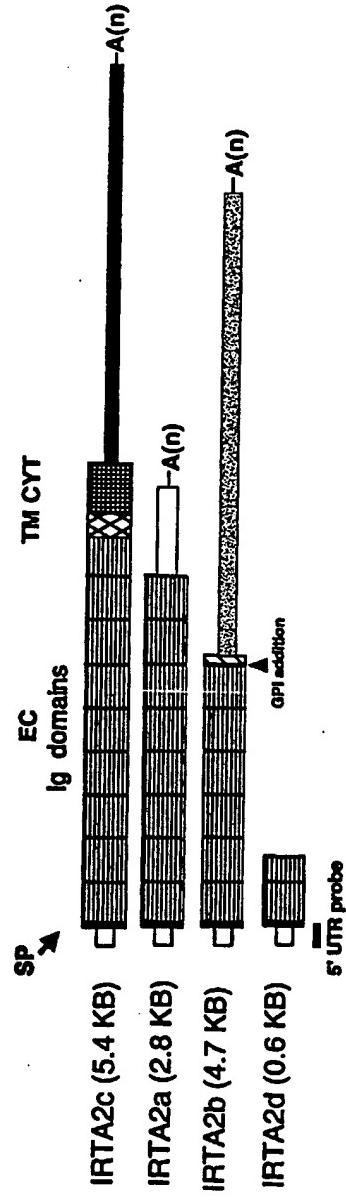


FIG. 9B



15/34

FIGURE 10

FIG. 10A

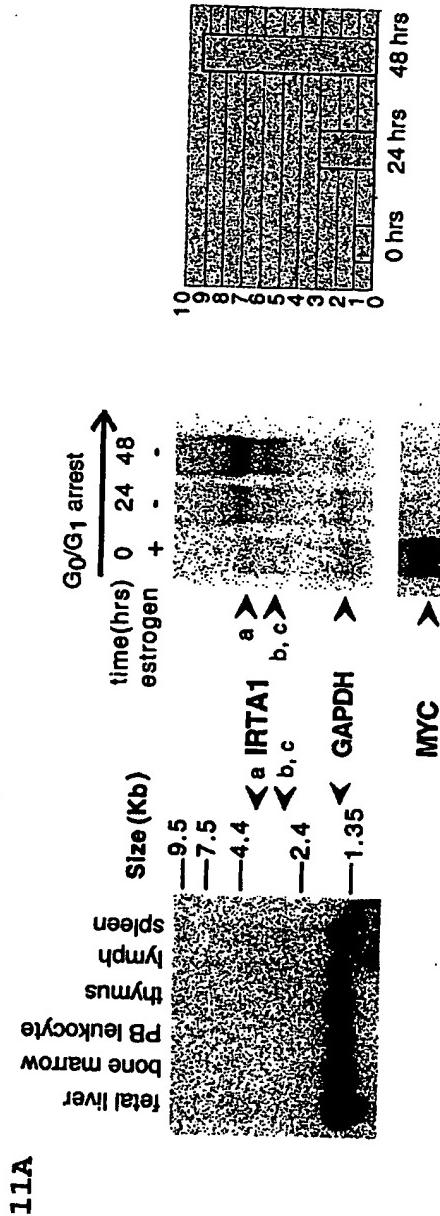
СИМВОЛЫ АРГАВА

FIG. 10B

<p><b>ITAM:</b></p> <p><b>IRTA1:</b></p> <p><b>IRTA2C:</b></p> <p><b>BGFI:</b></p>	<p><b>ITIM:</b></p> <p><b>IRTA1:</b></p> <p><b>IRTA2C:</b></p>	<p><b>ITAM:</b></p> <p><b>IRTA1:</b></p> <p><b>IRTA2C:</b></p>
<pre>XXXXXXXXX--DXX--XXXXXXX--XXXX</pre> <pre>SSSHSICPAQVELQSLTVVHPPKG-DLVYSEKOTTTGEEEBANTSRTLLEDKDVSVYXSEV</pre> <pre>DNKEPLNSDVQTEYQVSSAENSHK-----DLGKDKETVYXSEV</pre> <pre>DSDSQ---EPTXHNPVAWELOFVYT-----NANPDEGNTVYXSEV</pre> <pre>ASDOR---DLTEHKPVSNNHTODSN-----DPNNKNEVYXSEV</pre>	<pre>XXXXXXXXX</pre> <pre>SSXVXXL</pre>	<pre>XXXXXXXXX</pre> <pre>SSXVXXL</pre>
<pre>V</pre> <pre>V</pre>	<pre>V</pre>	<pre>V</pre>
<pre>Z</pre>	<pre>Z</pre>	<pre>Z</pre>
<pre>T</pre>	<pre>T</pre>	<pre>T</pre>

16/34

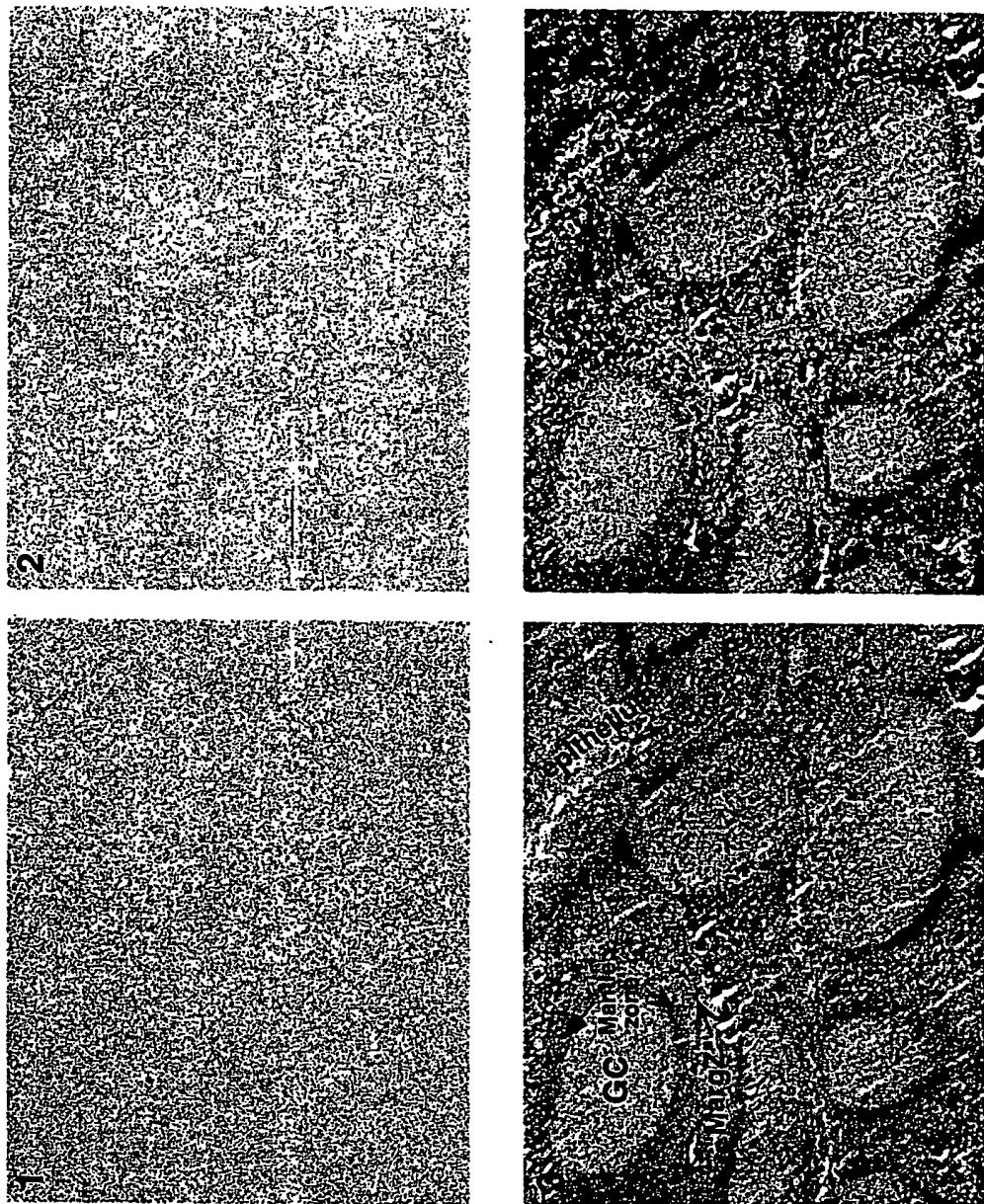
FIGURE 11A



17/34

FIGURE 11B1-B4

11B1-4



18/34

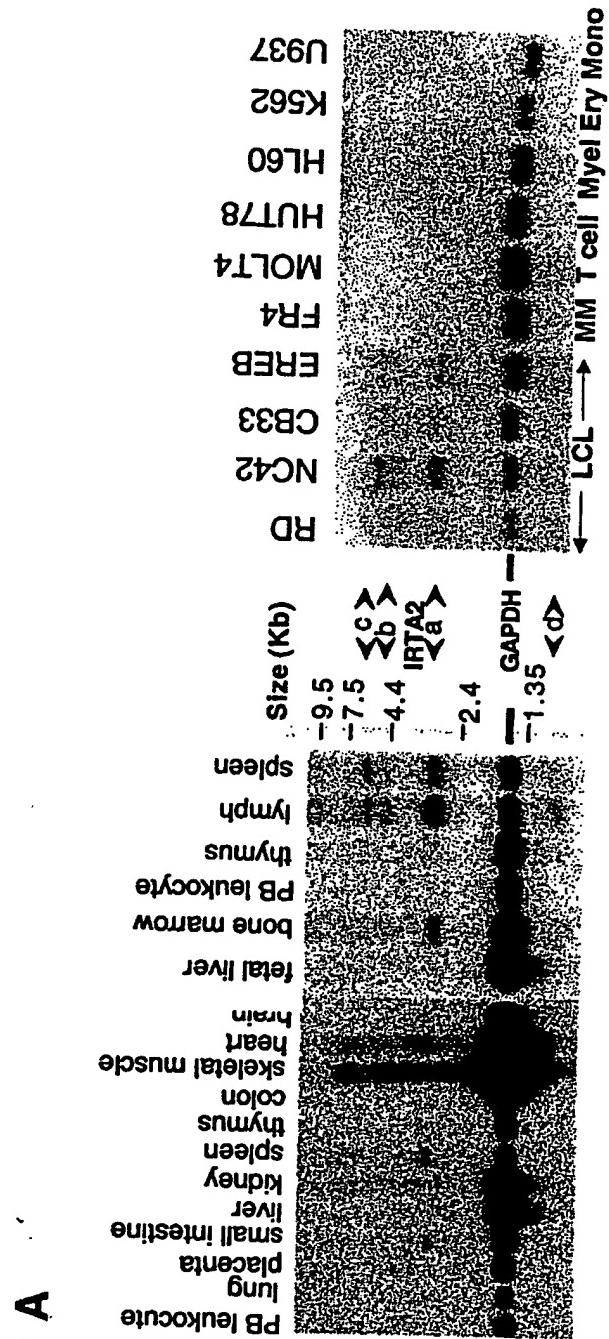
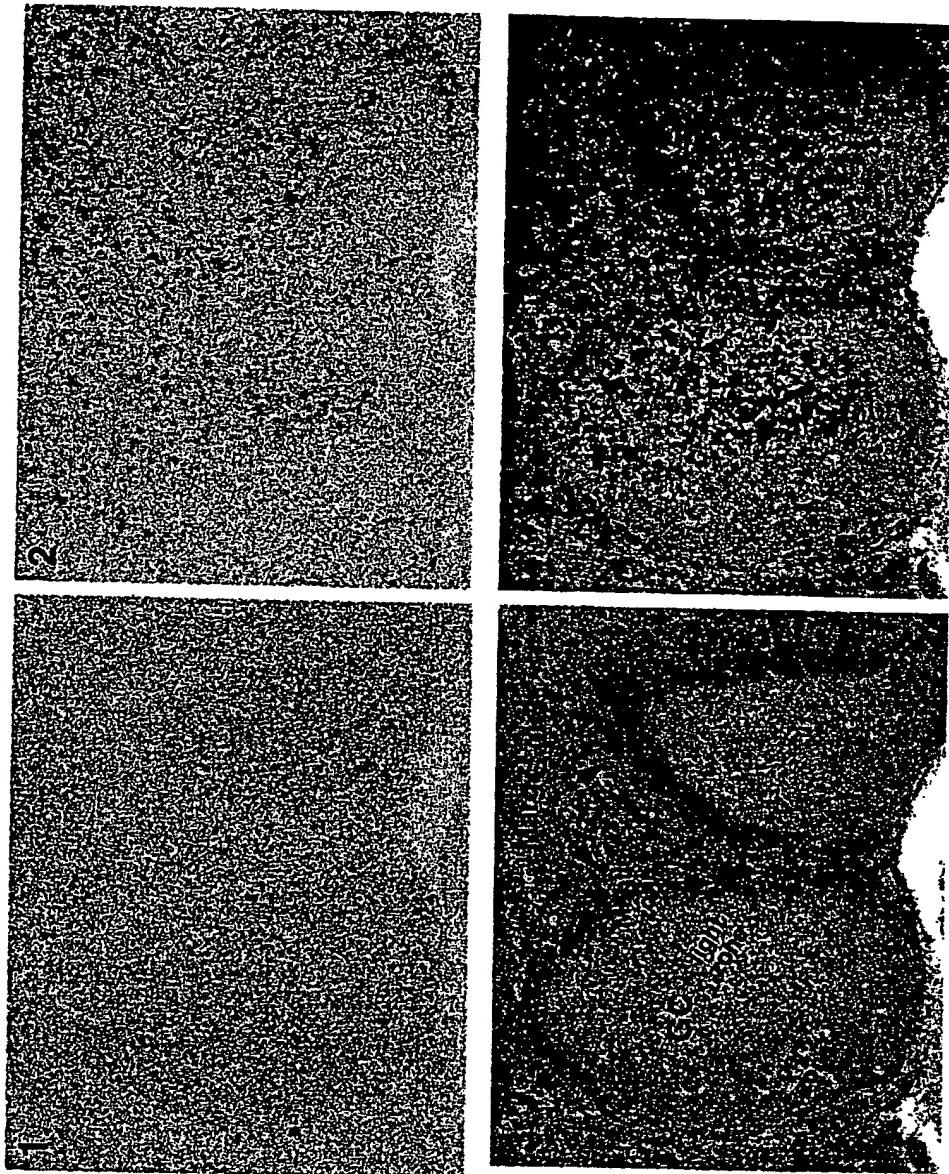


FIGURE 12A

19/34

FIGURE 12B1-B4

B



20/34

FIGURE 13

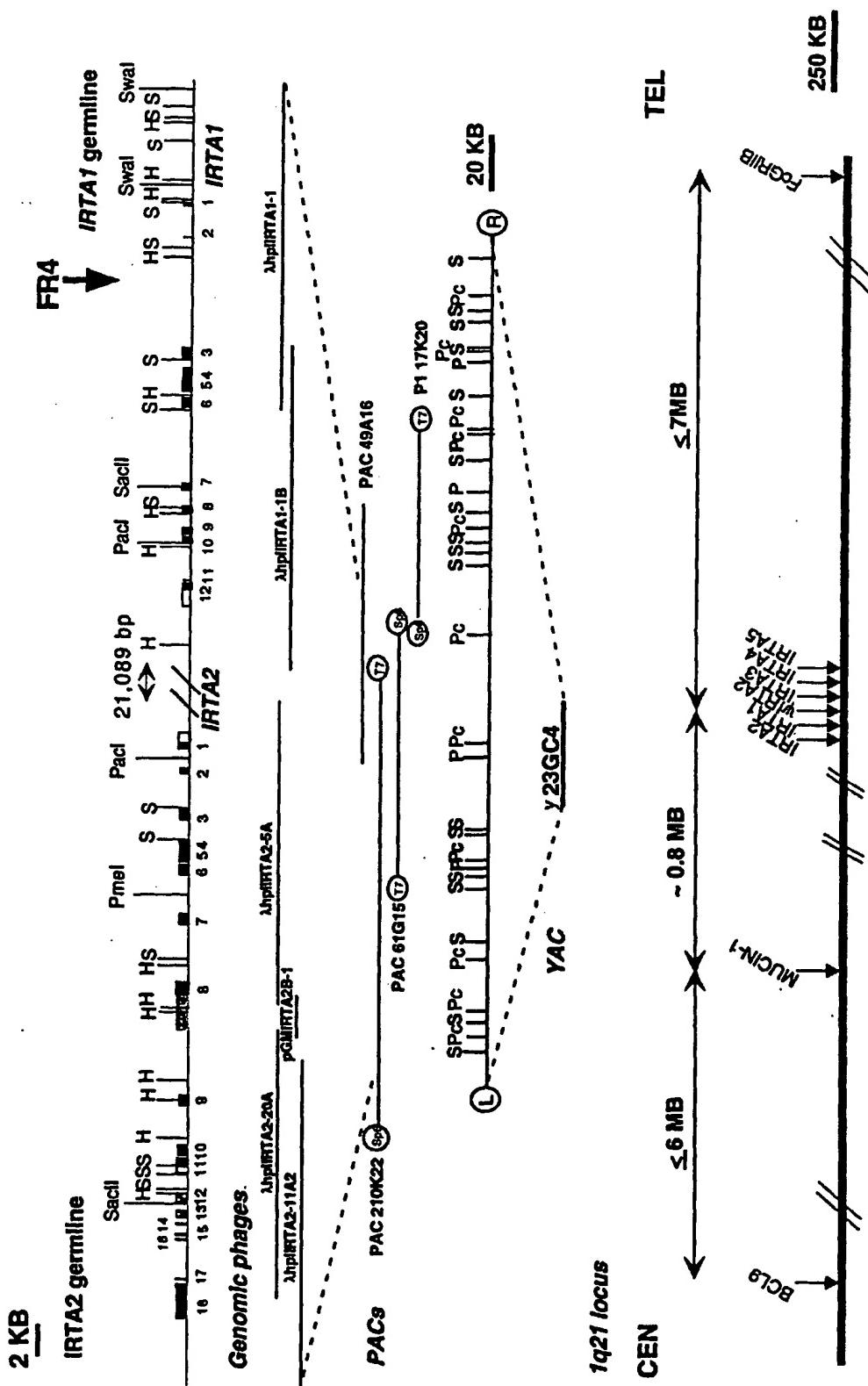


FIGURE 14

FIG. 14A

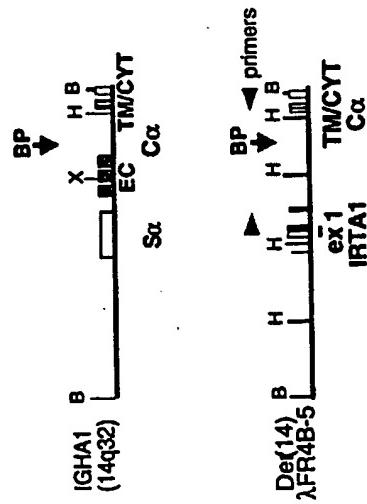


FIG. 14B

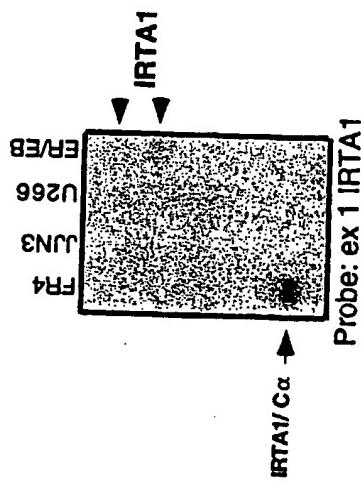


FIG. 14D

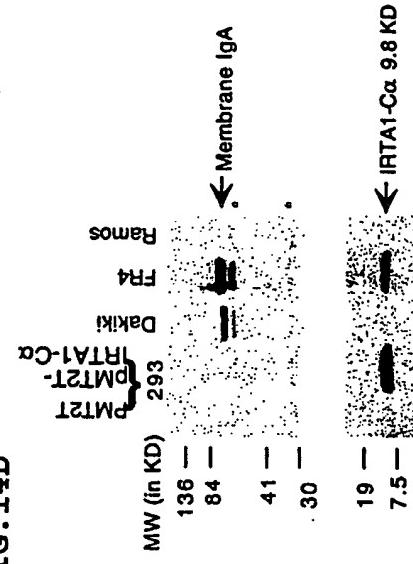


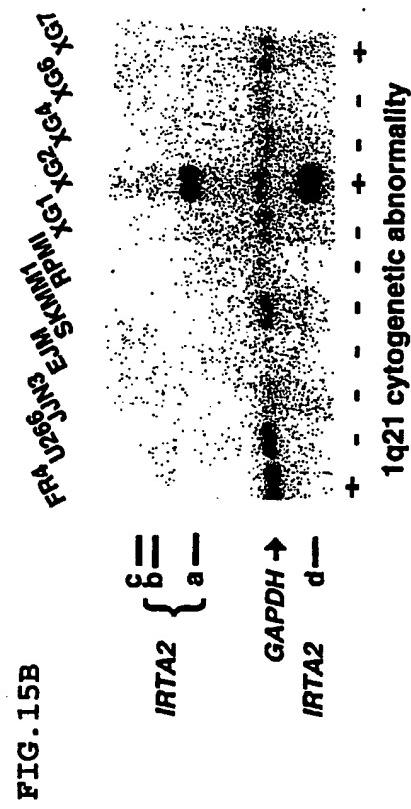
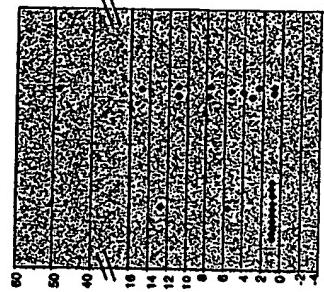
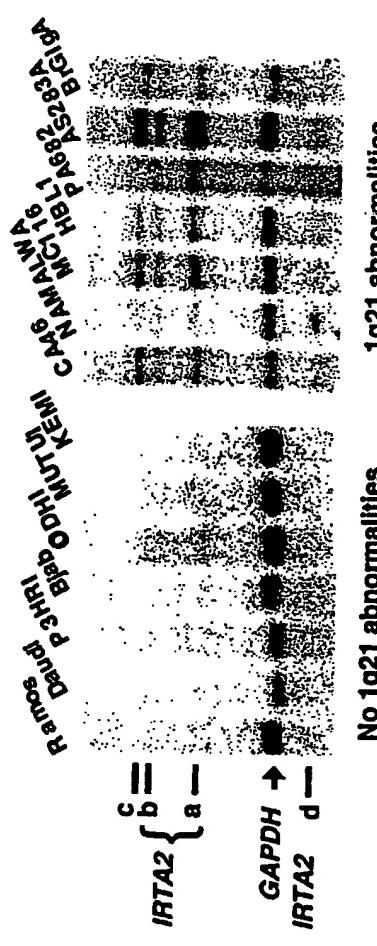
FIG. 14C



IP: anti-Cα (TM-specific)  
WB: anti-Cα (TM-specific)

## **IgG1 Heavy (top) and Light (bottom)**

FIGURE 15

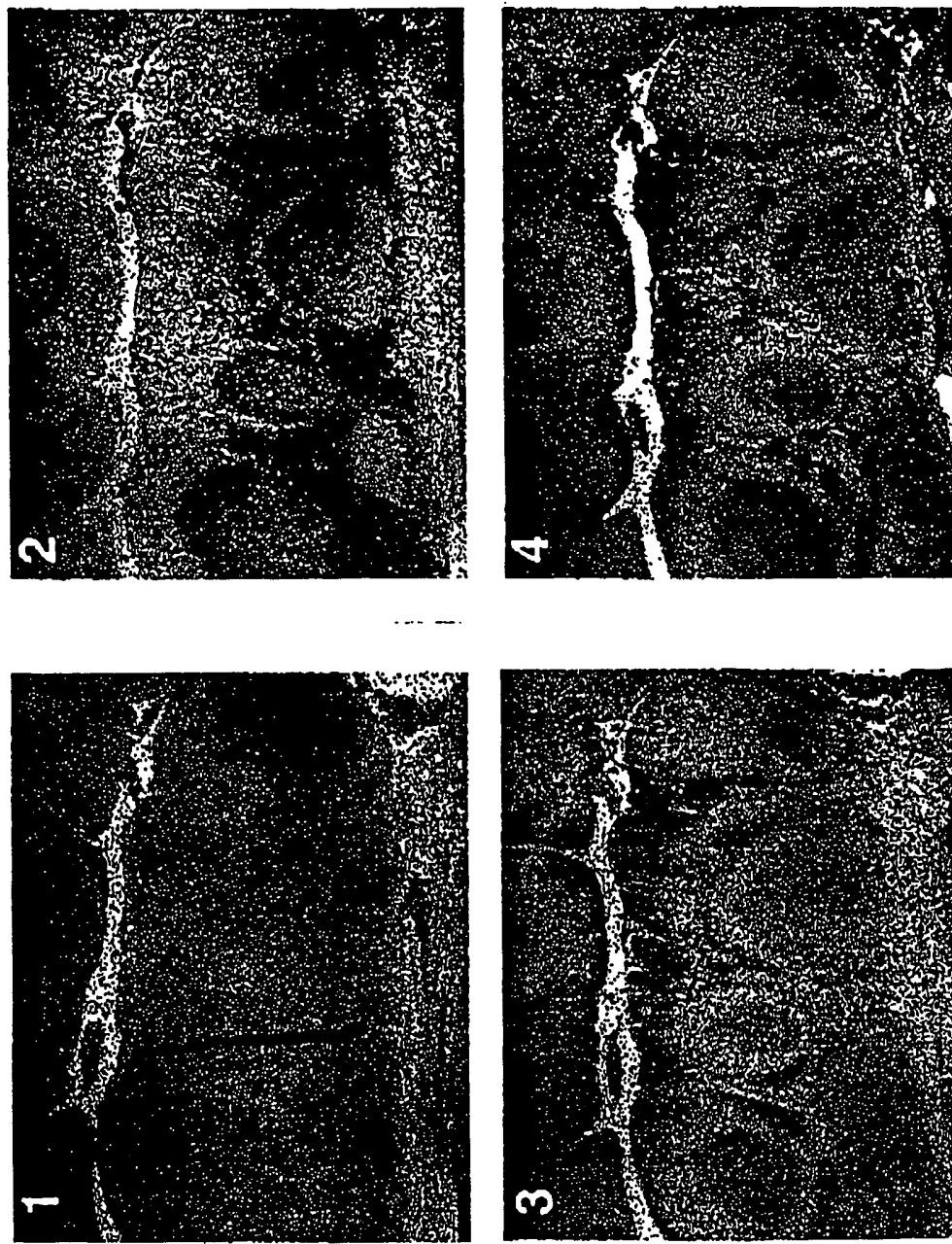


Cell line	<i>IRT A2</i>
Burkitt Lymphoma	2/12
Normal 1q21	
Abnormal 1q21	10/12
Multiple Myeloma	0/7
Normal 1q21	1/3
Abnormal 1q21	

Summary of *IRT A2* expression

23/34

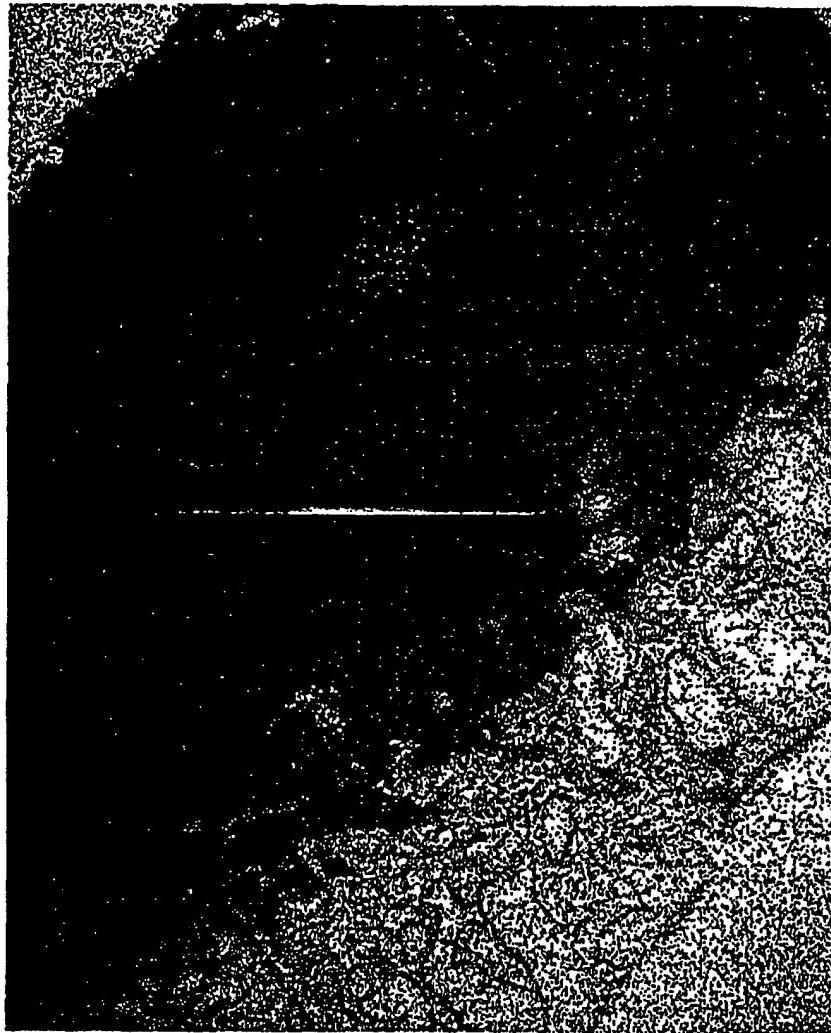
**FIGURE 16-1~16-4**  
**IRTA1 expression in normal lymphoid tissue**



24/34

FIGURE 17

**IRTA1 expression in a stomach Mucosa-Associated-Lymphoid Tissue B cell lymphoma**



25/34

**FIGURE 18A**

1	CTCAATCAGCTTATGCGAGAGAAGAGCTTACTGAGCTCACTGCTGGTCTGGTAGGCAAGTGTGCTTGGCAA L L L W A S 6
78	TCTGGGCTGACCTGGTTGTCTCCCTCAAGAACTCCCTCTCCAAACCCCTGGAGCAGGCCCTCCATGCTGCTGGGGCCCTCC L L A F A P V C G D S A A A H K P V I S V H P P W T 32
155	TTGCTGGCTTGTCCAGTCTGTGGACAATCTGCAGCTGCACACAAACCTGTGATTTCGTCATCCCATGAC T F F K G E R V T L T C N G F Q F Y A T E K T T W Y 58
232	CACATTCTCAAAGGAGAGAGACTCTGACTTGCAATGGATTCTGACTCTCATGCAACAGAAAACAATGGT H R H Y W G E K L T L T P G N T L E V R E S G L Y 83
309	ATCATCGGCACTACTGGGAGAAAAGTTGACCCCTGACCCCGAGAACACCCCTGGAGTTGGGAAATCTGACTCTAC R C Q A R G S P R S N P V R L L F S S D S L I L Q A 109
386	AGATGCCAGGCCCCGGCTCCCCACAGAACTAACCTGTCGGCTGCTCTTCAGACTCTTAATCTGAGGC P Y S V F E G D T L V L R C H R R K E K L T A V K 135
463	ACCATATTCTGTGTTGAGCTCACATTGGTTCTGAGATGCCACAGAAAAGGAAATTTGACTGCTGTGA Y T W N G S I S N K S W D L L I P Q P A S S N 160
540	AATAACTTGGAAATGAAAATCTTCTTCATTAATAAAAGCTGGGATCTCTTATCCCACAGCAAGTCAAT N N G N Y R C I G Y G D E N D V F R S N F K I I K I 186
617	ACAAATGGCAATTATCGATGCAATTGGATATGGAGATGAGAATGATGATTAGATCAAATTCAAAATTAAAT Q E L F P H P E L K A T D S Q P T E G N S V [REDACTED] C 212
694	TCAAGAACATTCCACATCCAGAGCTGAAAGCTACAGACTCTCAGCTACAGAGGGAAATTCTGTAACACCTGAGCT E T Q L P P E R S D T P L H F N F F R D G E V I L 237
771	GTGAAACACAGCTCTCCAGAGCGGTCAAGACACCCCACTTCACTTCAGAGATGGCGAGGTCACTCTG S D W S T Y P E L Q P L T V W R E N S G S Y W C G A 263
848	TCAGACTGGAGCACGTACCCGAACTCCAGCTCCAAACCGCTGGAGAGAAAACCTAGGATCTATTGGTGTGGTC E T V R G N I H K H S P S L Q I H V Q R I P V P S G V 289
925	TGAAACAGTGGGGTAACATCCAACAGCACAGTCCCTCGCTACAGATCATGTGAGCGGATCCCTGTCTGGGG L L E T Q P S G G Q A V E G E M L V L U C S V A E 314
1002	TGCTCTGGAGACCCAGCCCTCAGGGGCAAGGTGTTGAAGGGAGATGCTGCTCTGTCTCCCTGGCTGAA G T G D T F T S W H R E D M Q E S L G R K T Q R S L 340
1079	GGCACAGGGGATACCACATTCTCTGGCACCGAGAGGACATGCAGGAGAGTCGGGGAGAAAACCTAGCGTCCCT R A E L L P A I R Q S H A G G Y Y C T A D N S Y G 366
1156	GAGCAGAGCTGGAGCTCCCTGGCATCACAGAGGCCATGCAAGGGGATACTCTGACAGCACACAGCTACG P V Q S M V L [REDACTED] V R E T P G N R D G L V A A G 391
1233	GCCCTCTCAGAGCATGGTCTGAATGTCACTGTGAGAGAGACCCAGGCAACAGAGATGGCTTGTGCCCGGGGA A T G G L L S A L L A V A L F P W C H R R K S G 417
1310	GCCACTGGGGCTGCTCACTGCTCTCTGGCTGTGGCCCTCTGTTCACTGCTGGCTGGAGAGTCAGG V G F L G D E T R L P P A P G P G E S S H S I C P A 443
1387	AGTGGTTCTGGAGACGAAACCAAGGTCCCTCCGGCTCAGGCCAGGAGACTCTCCCATCTGCCCC Q V E L Q S L Y V D V H P K K G D L V X S E I Q T 468
1464	CCCAGCTGGAGCTCAGTGTATGTTATGATGACACCCCCAAAAGGGAGATTGGTATACTCTGAGATCCAGACT T Q L G E E E E A N T S R T L E D K D V S V V X S 494
1541	ACTAGCTGGAGAGAAGAGGAAGCTTAATACCTCCAGGACACTCTAGAGGATAAGGATGTCTAGTTGTCTACT E V K T Q H P D N S A G K I S S K D E E S * 515
1618	TGAGCTAAAGACACAAACCCAGATAACTCAGCTGGAAAATGACACTCTAGGATGAAGAAAGTAAAGAGAAATAAA 1695
1772	AGTACGGGAACGTCTACTCATGTGATTCTCCCTGTCAAAGTCCCAGGCCAGTGCACTCTGGGGCACCTG GAATGATCAACTATCCAGCTTCTAATCTCTCATGATGATGCTACTCCAGGAATACTCATGCTACT 1772
1849	CTGATGTTGGATGGAAATGGCCTCTGAAAGACTTCACTAAATGACCAGGATCCACAGTTAAGAGAAGCCCTGTAG 1926
2003	TATTTGCTGGGGCTGACCTAATGCACTTCTGGCTGTGGTTAGAGAAGGGGATAAAGAGAGAGAGGACTGT TATAAAGAACAGAACAAATTGGTGTGAATGGGATTTGCAAGAGATGAAAAAGACTGGGTGACCTGGATCTGCG 2080
2157	TIAATACATCAACCATGTCCTACTGCGAGAGACTACTGCACTAGTGTGTTACTCTGAGTGGCTGCACACGCA CTGTCACAAATGAAAGGCCCTCTACTCTGGCTGCAAGCTTACACTGTCAGGATTCTGAGATTAAGAAG 2234
2311	CCCACATGGAAATGGTTACAGAGAGGAATTAAAGGAGACATCACAGAGACTGGGAGATGCAAGCTCTAGGTG GCTTCCAAAAGCAAATGATAATTGTTAATGTCATTAGTGCACAAAGATTGCAACATTAGAGAAAAGACACAAA 2388
•	TATAAATTTAAAACCTAAAGTACCAACTCTCCAAACTAATTGAACTTTAAATTAGTATAAACTCATATAATAAA CTCTGCTTCTTAAATAAAAAAAAAAAAAA

## IRTA1 cDNA and protein sequence

26/34

FIGURE 18B-1

IRTA2A -----  
 IRTA2C -----  
 IRTA2B 1 CGGTGCAGTGTCCGTACTGTAAGATCAAGTCACCAACTGTTGGAAATTGAGGAACCTCTCTTGATCTCAGCCCTTG  
           M L L W V I L L V L A P V S G Q F A R T P R 22  
           81 GTGGTCCAGGTCTTCATGCTGCTGGGTGATATTACTGGTCCCTGGCTCTGTCAGTGACAGTTGCAAGGACACCCAG  
           P I F L Q P P W T T V F Q G E R V T L T C K G F R F 49  
 161 GCCCATTTTCTCCAGGCCATGGACCACAGTCTTCCAAGGAGAGAGACTGACCCCTCACTTGCAAGGGATTGCGCT  
           Y S P Q K T K W Y H R Y L G K E I L R E T P D N I L 75  
 241 TCTACTCACCAACAGAAAACAAAATGGTACCATGGTACCTTGGAAAGAAAATACTAAGAGAAACCCAGACAAATATCCTT  
           E V Q E S G E Y R C Q A Q G S P L S S P V H L D F S S 102  
 321 GAGGTCAAGGAATCTGGAGAGTACAGATGCCAGGGCTCCCTCTAGTAGGCCCTGTCAGTGGATTGGATTTCTTC  
           A S L I L Q A P L S V F E G D S V V L R C R A K A E V 129  
 401 AGCTTGGCTGATCTGCAAGTCACCTTCTGTGTTGAAGGAGACTCTGTGGTCTGAGGTGGCCAAAGGGAAAG  
           T L [REDACTED] I Y K N D N V L A F L N K R T D F H I P H 155  
 481 TAACACTGAAATATCTATTACAAGAAATGATAATGCTGGCATCTCTTAATAAGAAACTGACTCCATATTCTCAT  
           A C L K D N G A Y R C T G Y K E S C C P V S S N T V K 182  
 561 GCATGCTCAAGGACAATGGTCATGCTGACTGGATAAGAAAAGTGTGCCCCTGTTCTCCAAATACAGTCAA  
           I Q V Q E P F T R P V L R A S S F Q P I S G N P V T L 209  
 641 AATCCAAGTCAAGGCCATTACACGTCAGTGTGAGAGGCCAGTCCAGGGCCATCAGGGAAACCCAGTGACCC  
           T C E T Q L S L E R S D V P L R F R F F R D D Q T L 235  
 721 TGACCTGTGAGGCCAGCTCTCTAGAGAGGTCACTGCTCCGCTCCGGTCTTCAGAGATGACCAAGGCC  
           G L G W S L S P N F Q I T A M W S K D S G F Y W C K A 262  
 801 GGATTAGGCTGGAGTCTCTCCCGAATTCCAGATTACTGCCATGTGGAGTAAAGATCAGGGTTCTACTGGTGAAGGC  
           A T M P H S V I S D S P R S W I Q V O I P A S H P V L 289  
 881 AGCAACAATGCCCTCACGCCATCTGACAGGCCAGATCCTGGATAACGGTGCAGATCCCTGCATCTCATCTGTCC  
           T L S P E K A L N F E G T K V T L H C E T Q E D S L 315  
 961 TCACTCTGAGGCCCTGAAAAGGCTCTGAAATTGTGAGGAACCAAGGTGACACTCTACTGTGAAACCCAGGAAGATTCTG  
           R T L Y R F Y H E G V P L R H K S V R C E R G A S I S 342  
 1041 CGCACTTGTACAGGTTTATCATGAGGGTGTCCCCCTGAGGACAAGTCAGTCGCTGTGAAAGGGAGCATCCATCAG  
           F S L T T E N S G N Y Y C T A D N G L G A K P S K A V 369  
 1121 CTTCTCACTGACTACAGAGAATTAGGGAACTACTACTGACAGTCAGCTGACAATGCCCTGGCCAGGCCAGTAAGGCTG  
           S L S T V P V S H P V L [REDACTED] S P E D L I F E G A 395  
 1201 TGACCTCTCACTCACTGTGAGGCCAGAGGGTCACTCCCATCTCTGAGGACCTGATTGGAGGCC  
           K V T L H C E A Q R G S L P I L Y Q F H E D A A L E 422  
 1281 AAGGTGACACTTCACTGTGAAGGCCAGAGGGTCACTCCCATCTGAGGACCTGACATGAGGATGCTGCCCTGGA  
           R R S A N S A G G V A I S F S L T A E H S G N Y Y C T 449  
 1361 GCGTAGGTCGCCAACCTGCAAGGAGGTGCCATCAGCTCTCTGACTGAGCAGCATTAGGGAACTACTACTGCA  
           A D N G F G P O R S K A V S L S I T V P V S H P V L 475  
 1441 CAGCTGACAATGGCTTGGCCCCCAGGCCAGTAAGGCCGTGAGGCCCTCCATCTGCTCTGCTCATCTGTCCCT  
           T L S S A T L F E G A T V T L H C E V Q R G S P Q 502  
 1521 ACCCTCAGCTGCTGAGGCCCTGACTTTGAAGGGGCCACTGTGACACTCTACTGTGAGTCAGAGAGGGTCCCCACA  
           I L Y Q F Y H E D M P L W S S S T P S V G R V S F S F 529  
 1601 AATCTATACCAAGTTTATCATGAGGACATGCCCTGTGGAGCAGCTCAACACCCCTGTGGAAAGAGTGTCTCAGCT  
           S L T E G H S G N Y Y C T A D N G F G P Q R S E V V 555  
 1681 TCTCTGACTGAAGGACATTAGGGAAATTACTACTGCAAGCTGACAATGGCTTGGTCCCGAGGCCAGTGAGTGGTG  
           S L F P T V P V S R P I L T L R V P R A Q A V V G D L 582  
 2A, 2C1761 AGCCTTTGTCACTGTCAGTGTCTGCCACCTCTCACCTCACGGTCTCCAGGCCAGGCTGGTGGGGACCT 582  
 2B 1761 -----  
           GK C W V L A S H P P L A E F S L T H S F K 582  
           -----  
 2A, 2C1841 L E L H C E A P R G S P P I L Y W F Y H E D V T L G S 609  
 2B 1841 GCTGGAGCTTCACTGTGAGGCCAGAGGGCTCTCCCCAATCTGACTGGTTTATCATGAGGATGTCACCCCTGGGA  
           N L F A L S S F L P \* stop 592  
 2B 1841 AAATCTGTTGCACTGTCCAGTTCTCCCTTAATCAACCTTAATCCCCCTGTGGCTCTCTCAACTAACTAGCTGGG  
           S S A P S G G E A S F [REDACTED] L T A E H S G [REDACTED] C E 635  
 2A, 2C1921 GCAGCTCAGCCCCCTCTGGAGGAGAAGCTCTTCAACCTCTCTGACTGCAAGAACATCTGGAAACTACTCATGTGAG  
 2B 1921 GTTTCCGACTCATAGTCTGGCTAGCCAGCCCCCTAAACAGCTCAGTAGATCCCCAGCTTACCAAATGAATT  
           A N N G L V A Q H S D T I S L S V I V P V S R P I L T 662  
 2A, 2C2001 GCCAACATGGCTAGTGGCCCAGCACAGTGACACAAATCACTCAGTGTATAGTCCAGTATCTGTCACCCATCTCAC  
 2B 2001 TATTTATGTATTTCTCTCATCTGTCACAGTGGCAAGCTTACAGTGGCTTACAGTGGCAAGGCTCTCCCAA  
           F R A P R A Q A V V G D L L E L H C E A L R G S S P I 689  
 2A, 2C2081 CTTCAAGGGCTCCCAGGGCCAGGCTGTGCTGGGGACCTGCTGGAGCTTCACTGTGAGGCCCTGAGAGGCTCTCCCAA  
 2B 2081 TCTCTACTGACATTTACATTTACTAGCTACAAGCACAGTCTTATAGATAAAATTGGTCAAGACCTTAAATTCTCCA

27/34

## FIGURE 18B-2

L Y W F Y H E D V T L G K I S A P S . G G G A S F N E D O  
 2A, 2C2161 TCCCTGACTGGTTTATCATGAAGATGTCACTGGGTAAGATCTAGCCCCCTCTGGAGGAGGGCTCCTTCAGCTCC  
 2B 2161 AAGGATTCCAATCTTATGGTAGATTGGAGAAAGCTGCTGGTAACAAAGGGGAAATGGCTCTAGGAACCAACTCC 715

L T T E H S G I Y S C E A D N G L E A Q R S E M V T 742  
 2A, 2C2241 TCTCTGACTACAGAACATCTGGAACTACTCTGTGAGGAGCACAACTGGCTGGAGGCCAGCGCACTGAGATGGTGAC  
 2B 2241 TCAAACCTCTGGAGTTTATGATCCCTGTTCTAACCTGCTAAAATCAGTATCATTTATTGTTAAAAAA

L K V A V P V S R P V L T L R A P G T H A A V G D L L 769  
 2C 2321 ACTGAAAGTTGCAGTCCGGTGTCTGCCCGGTCTCACCCCTCAGGGCTCCGGGACCCATGCTGGGTGGGGACCTGC  
 ----- G E W A L P T S S T S E N \* 759  
 2A 2321 -----GGTAGTGGGCCCTGCCACCAGCAGCACATCTGAGAACCTGACTGTGCTGTCTCCCTGCAGCTGA  
 2B 2321 ACTATTGTAAGTATGACATCACATTCAAGAAACGTGTGCAAATTGATGTGTAACGATTGGTGTCTTTAGGAGCTAA

E L H C E A L R G S P L I L Y R F F H E D V T L G 795  
 2C 2401 TGGAGCTTCACTGTGAGGCCCTGAGGGCTCTCCCTGATCTGTACCGGTTTTTCATGAGGATGTCACTCTAGGAAAT  
 2A 2401 AAATGGAGCCACAGAGCTCTCAGGGCTGTTGTGTCAGCACACTCTCTGCCCTGCAGAACCTCCCTGTG  
 2B 2401 GTTGCCTCTGTTTACTGAACTTGTGTTATAGAAACTGGGAAAGTTACTTCTTCACTGAGAGAAATGGTA

S P S G G A S L L T A E H S G C E A D 822  
 2C 2481 AGGTGTCCTCCCTCTGGAGGAGCGCTCTAAACCTCTCTGACTGAGCACACTGAAACTACTCTGTGAGGCCGA  
 2A 2481 AAAGTCTCGGATCTTGTGTTAGCTGTCAGGAACTCTGATGTGTTCCAGCAGTCTCTGAAGATGATAAGAACCTC  
 2B 2481 TGATAGAAAATCTTGAGGCTGATGTGTCAGACATGCCCTAGCATACTGAGTAAAGAGGTTATTTAAAATGT

N G L G A Q R S E T V T L Y I T G L T A S G P F A 849  
 2C 2561 CAATGGCTCGGGGCCAGCGAGCTGAGGAGACTGACACTTTATACAGGGCTCACCGCGAACAGAACGGCCCTTTG  
 2A 2561 ACTAAAATGCAAAATAGACTTTTAAAGACATAACTATTCTGACTGAATTTAACATGAAATTTAACATGAAATGAAACCAAGA  
 2B 2561 GAATGTTCTGAGACTACTCAGAACAGCTACTAGGAAGCTCTAGACTCACTCATCTGCATCCATTAC

T G V A G G L L S I A G L A A G A L L Y C W L S R 875  
 2C 2641 CCACAGGAGTCGGGGCCCTGCTCACGATAGCAGGCTCTGCTGGGGGACTGCTGCTACTGCTGGCTCTCGAGA  
 2A 2641 ATTCTGAGCATATGTTCTCTGGCTAGAAAGGATAAGCTGTTCTGTCAGGATTTCTCATTTGACTCTTAAGAA  
 2B 2641 TATTTTATCCATGTTTACTTCTCTCATATTAGCAGCATCTTAAGCTCTTATTTCTGTTCTGACTGTCA

K A G R K P A S D P A R S P ' S D S D S Q E P T Y H N V 902  
 2C 2721 AAAGCAGGGAGAAACCTGCCCCTGACCCGCCAGGAGCCCTCAGACTCGACTCCAAAGGCCACCTATCACATGT  
 2A 2721 GCCTCTACTCTTGAGCTCTTCAATTACTGGGGATGAAATGTCCTACATTCAACATTAAATCTATGTTAACGA  
 2B 2721 CCCTTAATGCCAGTAGAAATGTAAGCTCATGAGAACAGAACCTGCATCCATCTGGCTCTCACACATCCCTGTGCCTACT

P A W E E L Q P V Y T N A N P R G E N V V Y S E V R I 929  
 2C 2801 ACCAGCTGGGAGAGCTGCAACCAGTACACTAATGCAAACTCTAGAGGAGAAATGTGGTTACTCAGAAAGTACCGA  
 2A 2801 AAAA  
 2B 2801 CAGTGGTGGCACACAGTAGGTCTCAGTCACATTGTAATTAGTGACAGATGATATGACAAGATGATAAGAGGGGA

I Q E K K K H A V A S D P R H L R N K G S P I I Y S 955  
 2C 2881 TCATCCAAGAGAAAACATGCCAGTGGCTCTGACCCCTGAGGAACTCTCAGGAACAAAGGGTTCCCTATCATCTACTCT  
 2B 2881 TTTAAAAAAATCATCTAGCAAGGCCAAAGGAAAAAAACAAAGCTTTAGAAATGAAATACCAATTGAAAGCAGTA

E V K V A S T P V S G S L F L A S S A P H R \* stop 977  
 2C 2961 GGAGTTAGGTGGCTCAACCCGGTTCCGGATCCCTGTTCTGGCTCTCAGGCTCTCACAGATGAGTCCACACGTC  
 2B 2961 AGAATAGATTGGATATCTTGAAAACCTTAATTGAAATGAAAGACCAATTGAGAAAACAATACAGAAATGCAAGTAGAA

2C 3041 TCTCCAACCTGTTCTGACCCCTGACCCCAAAGTTCCCTGGGGAGAACAGCAGCTGAGTGGAAAGATTAGGCT  
 2B 3041 AGATACAGAAAATAGCAAAAGTTATAATGAAATCAGACAATGGATTGTCATCTGAGTTGATAATGTTGAGATAAAGAAA

2C 3121 GCCCCAGACCATATCTACTGGCTTGTTCACATGTCCTCATTCTCAGTCAGCAGAACATGCAAGGGCTCTGGACTG  
 2B 3121 AATGGGAGCCCTCAGAAAATTGAAACCGAGAGTAAAGTCAACTAAAAATGAGTAGAAATTGTTGGAGATAAAGAAA

2C 3201 TCACCTGTTCCCTAGTAAAGCCCTGACTGGCAGGTTTTAATCCAGTGGCAAGGTGCTCCACTCCAGGGCCAGCAC  
 2B 3201 ACTTGAATATGAGATCAGAACATATGTTGATGACGTTAGCTGACTTTGAGGTTAAAATATATATGTCCTATGAT

2C 3281 ATCTCCGGATTCTTGTGGGCTTCAGCTGTTGAGTACTGCTCTCATCACACCCCCACAGGGGGTC  
 2B 3281 TATGGGAAAAAAAGCAGTCGCTCAGAAAGAAAACATCAAGTTAGTCTAGACTTTGAGCTGACTCAGTACCAAGAG

3361 TTACCCACAAAGGGAGATGGGCTTCAGGAGATGCCGGCTGGCTAACAGCTCAGGTGCTCTAAACTCCGACACAG  
 3441 AGTTCTGTTGGGGATGTCATTTCTCAATTGTCATCACCTGGGGCTACTGCACTGTCAGTGTGCTGCCAAATGGGACAG  
 3521 CACACAGCTGTGACATGGGACATGTGATGGGTCTCCCTCAGGGGGCTGCAATTCAACTCTCCACCTGCTCCTAAACT  
 3601 CTAGGTGGCACTTGACACCAAGGTAACCTCTCTCTGCTCATGTCAGTGTCTACCTGCCCAAGTAAGTGGCTTCA  
 3681 TACACCAAGTCCCGAAGTTCTCCCTCTCAACAGAGTAACCCAGCAAGTCAAGGCCAGGAGGACCAAGGGGTGAGACA  
 3761 GAACACATACTGGAACACAGGAGGTGCTCAATTACTATTGACTGACTGACTGAATGAATGAATGAGGAAGAAAAC  
 3841 TGCGGTAATCAAACCTGGCATAAAATCCAGTCAGCAAGAAGAAGGGCTCTCACACTTTAGTGTGCTTGTGAGGAC  
 3921 GAAGAGAAGGAGCTGGATGAAAGAAAATGTTCAAGCAAGAAGAAGGGCTCTCACACTTTAGTGTGCTTGTGAGGAC  
 4001 GAGGATCTGAAAATACAGATACTGATTCACTGAGCTGCTGAGACTGCCATTCTAACATGTTCCAGGG

28/34

## FIGURE 18B-3

4081 ATGCTGATGCTGCTGGCCCTGGGACTGCACTGCATGCATGTGAAGCCCTATAGGTCTCAGCAGAGGCCCATGGAGAGGGA  
4161 ATGCTGGCTCTGGCTGCCAGGGCCAACTCGGTTACACGGATCTGCTGCCCTGGCCAGCCTTGGCCACAGCAC  
4241 CACCAACTGCTGTTCTGAGAGAGCTCTCTGACATGGATGTTCTGCTTATGATATATAACCTGAGCCCTCTCC  
4321 TGCCACTATCTTGTTCCCCACCTCAGGCCCTCACACTTCCCATGAAAAGGGTGAATGTATATAACCTGAGCCCTCTCC  
4401 ATTCAAGAGTTGTTCTCCCACATCTGAGCAATGGGATGTTCTGCTTATGATATCATCACATCTTATCTGATC  
4481 TTGCTCCCAAGTGGATTGACAGTGACTTTAAGCCCCACGGCCCTGAATAAAATCCTTCAAGGGCATGGAAAGC  
4561 TCACTCCACCTGAACCATGGCTTTCATGCTTCAAGTGTCAGGGCCTTGCCCAAGATAGACAGGGCTGACTCTGCTGCC  
4641 CAACCTTCAGGGAGAACCCAGACACCTGAGACAGGAGCTGTATGCAGCCCAGTGAGCTTGAGAGACAAGCTG  
4721 GAGGCATTGTCATCACTACAGATGCAACTAAATAGACGTGGAGCAAGAGAAATGCATTCCACCGAGGCCCTTT  
4801 TTAGGCCTACTTGAAAGTCAGAACAGCACGAGCAAGCATAGGCTCAGGATTAAGAAAAAAATCTGCTCACAGTCTGTT  
4881 CTGGAGGTACATCACCAAAAGCTACGCCCTATGCAGTTCTGAGAAGGTGGAGGCACCAGGCTCAAAGAGGAAATT  
4961 TAGAATTCTCATGGAGAGTAAGCTACCCCCATCCAGAATGATAACTGCACAGTGGAGAACAAACTCCACCTAAT  
5041 GTGGCTGGACCCCATCAGTCTGTTGAAGGCCCTGAATGTAACAAAAGGGTTATTCTCTCAAGTAAGGGGAACCT  
5121 GCTTGGGCTGGGACATAAGTTCTGCTTCAAGACGCAAACGTGAAAATGGCTCTCTGGGCTTGACCTTGCTGGC  
5201 ATATGGACTGAAAGAAACTATGCTATTGGATCTCTGGATCTCCAGCTGACTGCAGATCTTGAGATATGTCAGCCT  
5281 CTACAGTCACAAGAGCTAATTCTAATAAACCAATCTTC

29/34

## FIGURE 18C-1

1 AGTGAAGGGGTTCCCATATGAAAAATACAGAAAGAATTATTAATACTA  
 52 GCAAATACAACACTGTGATTTCTAGAGAACCCACGCACAGTCMTGGAGAC  
 103 ATTACTCTGAGAGACTGCAGCTGATGGAAGATGAGCCCCAACATCTAAA  
 154 ATGTATCACTACCGGGATTGAGATACAAACAGCATTTAGGAAGGTCTCATC  
 205 TGAGTAGCAGCTTCTGCCCTCTCTGGAGATAAGTCGGGCTTTGGTG  
 256 AGACAGACTTCCCACCCCTCTGCCCGGGTGCCCATGCTCTGTGGCT  
 1 M L L W L  
 307 GCTGCTGCTGATCTGACTCTGGAAAGAGAACAAATCAGGGTGGCCCCAAA  
 6 L L L I L T P G R E Q S G V A P K  
 358 AGCTGTACTTCTCTCAATCTCCATGGTCCACAGCCTCAAAGGAGAAA  
 23 A V L L N P P W S T A F K G E K  
 409 AGTGGCTCTCATATGCAGCAGCATATCACATTCCCTAGGCCAGGGAGACAC  
 450 V A L I C S S I S H S L A Q G D T  
 460 ATATTGGTATCACGATGAGAAGTGTGAAATAAAACATGACAAGATCCA  
 57 Y W Y H D E K L L K I K H D K I Q  
 511 AATTACAGAGCTGGAAATTACCAATGTAAGACCCGAGGATCCTCCCTAG  
 74 I T E P G N Y Q C K T R G S S L S  
 562 TGATGCCGTGATGTGAAATTTCACCTGACTGGCTGATCTGCAGGCTT  
 91 D A V H V E F S P D W L I L Q A L  
 613 ACATCCTGTTGAAAGGAGACAATGTCATTCTGAGATGTCAGGGAAAGA  
 108 H P V F E G D N V I L R C Q G K D  
 664 CAACAAAACACTCATCAAAGGTTACTACAAGGATGGAAAACAGCTCC  
 125 N K N T H Q K V Y Y K D G K Q L P  
 715 TAATAGTATAATTAGAGAAGATCACAGTGAATTCACTGTCAGGGATAA  
 142 N S Y N L E K I T V N S V S R D N  
 766 TAGCAAATATCATGTAUTGCTTATAGGAAGTTTACATACTTGACATTGA  
 159 S K Y H C T A Y R K F Y I L D I E  
 817 AGTAACCTCAAAACCCCTAAATATCCAAGTTCAAGAGCTGTTCTACATCC  
 176 V T S K P L N I Q V Q E L F L H P  
 868 TGTCGTAGAGCCAGCTCTCCACGCCATAGAGGGAGTCCCATGACCC  
 193 V L R A S S S T P I E G S P M T L  
 919 GACCTGTGAGACCCAGCTCTCCACAGGGCCAGATGTCAGCTGCAATT  
 210 T C E T Q L S P Q R P D V Q L Q F  
 970 CTCCCTCTCAGAGATAGCCAGACCCCTGGATTGGCTGGAGCAGGTCCCC  
 227 S L F R D S Q T L G L G W S R S P  
 1021 CAGACTCCAGATCCCTGCCATGTCAGACTGAAAGACTCAGGGTCTTACTGGT  
 244 R L Q I P A M W T E D S G S Y W C  
 1072 TGAGGTGGAGACAGTGTACTCACAGCATCAAAAAAGGAGCCTGAGATCTCA  
 261 E V E T V T H S I K K R S L R S Q  
 1123 GATACTGTACAGAGACTCCCTGTCTATGTGAATCTAGAGATCCGGC  
 278 I R V Q R V P V S N V N L E E I R P  
 1174 CACCGGAGGGCACCCTGATTAAGGAGAAAATATGGCTTTATTGCTCAGT  
 295 T G G Q L I E G E N M V L I C S V  
 1225 AGCCCAGGGTTCAAGGACTGTACATCTCTGGCACAAAAGAAGGAAGAGT  
 312 A Q G S G T V T F S W H K E G R V  
 1276 AAGAAGCCTGGTAGAAAGACCCAGCCTCCCTGTGGCAGAGCTGCATGT  
 329 R S L G R K T Q R S L L A E L H V  
 1327 TCTCACCGTGAAGGAGAGTGTGAGGGAGATACTACTGTGAGCTGATAA  
 346 L T V K E S D A G R Y Y C A A D N  
 1378 CGTTCACAGCCCCATCTCAGCACGTGGATTGAGTCACCGTGAGAATTCC  
 363 V H S P I L S T W I R V T V R I P  
 1429 GGATCTCACCCTGTCCTCACCTCAGGGCTCCAGGGCCACACTGTGGT  
 380 V S H P V L T F R A P R A H T V V  
 1480 GGGGGACCTGCTGGAGCTTCACTGTGAGTCCCTGAGAGGCTCTCCCCGAT  
 397 G D L L E H C E S L R G S P P I  
 1531 CCTGTACCGATTATCATGAGGATGTACCCCTGGGGAACAGCTCAGCCCC  
 414 L Y R F Y H E D V T L G N S S A P  
 1582 CCTGGAGGAGGAGCCTCTCAACCTCTCTGACTGAGAACATTCTGG  
 431 S G G G A S F N L S L T A E H S G  
 1633 AAACACTCTGTGATGCCAGAACATGGCTGGGGCCAGCACAGTCATGG  
 448 N Y S C D A D N G L G A Q H S H G  
 1684 ACTGAGTCTCAGGGTACAGATTCGGTGTCTGCCCTCTCACCCCTAG  
 465 V S L R V T V P V S R P V L T L R  
 1735 GCCTCCGGGGCCAGGCTGCTGGTGGGGACCTGCTGGAGCTTCACTGTGA  
 483 A P G A Q A V V G D L L E L H C E  
 1786 GTCCCTGAGAGGCTCTCCGATCTGTACTGGTTTATCACGGAGGATGA  
 499 S L R G S P I L Y W F Y H E D D  
 1837 CACCTGGGGAACTCTGGCCCACTCTGAGGAGGGCATCTTCACCT  
 516 T L G N I S A H S G G G A S F N L  
 1888 CTCTCTGACTACAGAACATTCTGGAAACTACTCATGTGAGGCTGACAATGG  
 533 S L T T E H S G N Y S C E A D N G

30/34

## FIGURE 18C-2

1939 CCTGGGGGCCACCAAGTAAAGTGGTGACACTCAATGTTACAGGAACCTC  
550 L G A Q H S K V V T L N V T G T S  
1990 CAGGAACAGAACAGGCCCTAACCCCTGCCGGAATCACGGGGCTGGTGCAG  
567 R N R T G L T A A G I T G L V L S  
2041 CATCCTCGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG  
584 I L V L A A A A A L L H Y A R A R  
2092 AAGGAAACCAGGAGACTTCTGCCACTGGAACATCTAGTCACAGTCCTAG  
601 R K P G G L S A T G T S S H S P S  
2143 TGAGTGTCAAGGAGCCTTCCTCGCCAGGCCCTCCAGGATAGACCCCTCAAGA  
618 E C Q P S S R P S R I D P Q E  
2194 GCCCACTCACTCTAAACCACTAGCCCCAATGGAGCTGGAGCCAATGTACAG  
635 P T H S K P L A P M E L E P M Y S  
2245 CAATGTAATCCTGGAGATAGCAACCCGATTATTCCAGATCTGGAGCAT  
652 N V N P G D S N P I Y S Q I W S I  
2296 CCAGCATACAAAAGAAAATCAGCTAATTGTCATGATGCATCAAGAGCA  
669 Q H T K E N S A N C P M H Q E H  
2347 TGAGGAACCTACAGTCCTCTATTCAAGACTGAGAAGACACACCCAGACGA  
686 E E L T V L Y S E L K K T H P D D  
2398 CTCTGCAGGGGGAGCTAGCAGCAGAGCAGGGCCATGAAGAAGATGATGA  
703 S A G E A S S R G R A H E E D D E  
2449 AGAAAACATGAGAATGTACCACTGTTATTACTGGCTCAGACCACTAGCC  
720 E N Y E N V P R V L L A S D H  
2500 CCTTACCCAGAGTGGCCCACAGGAAACAGCCTGCACCATTTTTCTGT  
2551 TCTCTCCAACCACACATCATCCATCTCCAGACTCTGCCTCTACGAGGC  
2602 TGGGCTGCAGGGTAATGTGAGGCTGAGCAAAGGCTGCAAATCTCCCTGT  
2653 GCCTGATCTGTGTGTTCCCCAGGAAGAGAGCAGGGCAGGCCCTCTGAGCAAGCA  
2704 CTGTGTTATTTTACAGTGAGACAGTGCCAGGCAGGAGGGCCCTCAGC  
2755 TCCTAGGGCTGTCGAATAGAGGAGAGAGAATGGCTAGCCAGGGTAA  
2805 CAAGGGCAACATGACCAATTGATCAAGTGATGATGAAAGCTGTTAAT  
2857 GTCTCTCTGTATAAACAAATTGCTCAAATATTGTTGTTCCCTTTTGT  
2908 GTGGCTGGTAGTGGCATTGCTGATGTTGGTGTATGCTGATCCTTGC  
2959 TACCATATTGGG

31/34

## FIGURE 18D-1

1 TGGTACCAAGAGTACATCTCTTTCAAATAGCTGGATTAGGTCCATGC  
 1 M L  
 52 TGCTGTGGTCATTGCTGGCATCTTGATGCAGTCAGTGAACAGGCAGATT  
 19 L W S L L V I F D A V T E Q A D S  
 103 CGCTGACCCCTGGGCCCTCTCTGCTCGAAGGAGACAGCATCGTC  
 36 L T L V A P S S V F E G D S I V L  
 154 TGAAATGCCAGGGAGAACAGAACAGTGGAAAATCAGAAGATGGCTTACCCATA  
 53 K C Q G E Q N W K I Q K M A Y H K  
 205 AGGATAACAAAGAGTTATCTGTTTCAA AAAATTCTCAGATTCCCTATCC  
 70 D N K E L S V F K K F S D F L I Q  
 256 AAAGTGCAGTTAAGTGACAGTGGTAACCTTCTGTAGTACCCAAGGAC  
 87 S A V L S D S G N Y F C S T K G Q  
 307 AACTCTTCTGGATAAAACTCAAAATATAGTAAAGATAAAAGTCCAAG  
 104 L F L W D K T S N I V K I K V Q E  
 358 AGCTTTCAACGCTCTGTGACTGCCAGCTCCAGCCATCGAAG  
 121 L F Q R P V L T A S S F Q P I E G  
 409 GGGGTCAGTGGCTGAANTGTGAGACCCGGCTCTCCACAGAGGTTGG  
 138 G P V S L K C E T R L S P Q R L D  
 460 ATGTTCAACTCCAGTCTGCTCTTCAGAGAAAACCAGGTCTGGGTCAAG  
 155 V O L Q F C F F R E N O V L G S G  
 511 GCTGGAGCAGCTCCGGAGCTCCAGATTCTGGCTGTGGAGTGAAGACA  
 172 W S S S P E L Q I S A V W S E D T  
 562 CAGGGCTTACTGGTCAAGGCAGAACCGGTGACTCACAGGATCAGAAAC  
 189 G S Y W C K A E T V T H R I R K Q  
 613 AGAGCTCCAATCCCAGATTACGTGCGAGAATCCCCATCTAAATGTAA  
 206 S L Q S Q I H V Q R I F I S N V S  
 664 CCTGGAGATCGGGCCCCGGGGACAGGTGACTGAAGGACAAAAGTGA  
 223 L E I R A P G G Q V T E G Q K L I  
 715 TCCTGCTCTGCTCAGTGGCTGGGGTACAGGAATGTCACATTCTCTGGT  
 240 L L C S V A G G T G N V T F S W Y  
 766 ACAGAGAGGCCAACAGAACAGTATGGAAAGAAAACCCACCGTCCCCGT  
 257 R E A T G T S M G K K T Q R S L S  
 817 CAGCAGAGCTGGAGATCCCACCTGTGAAAGAGATGATGCCGCAAAATATT  
 274 A E L E I P A V K E S D A G K Y Y  
 868 ACTGTAGAGCTGACACGGCATGTGCTATCCAGAGCAAGGTGGTGAATA  
 291 C R A D N G H V P I Q S K V V N I  
 919 TCCCTGTGAGAATTCCAGTGTCTCGCCCTGTCTCACCCCTCAGGTCTCC  
 308 P V R I P V S R P V L T L R S P G  
 970 GGGCCAGGCTGCAGTGGGGGACCTGCTGGAGCTTCACTGTGAGGCCCTGA  
 325 A Q A A V G D L L E L H C E A L R  
 1021 GAGGTCTCCCCAATCTGTACCAATTATCATGAGGATGTCAACCTTC  
 342 G S P I L Y Q F Y H E D V T L G  
 1072 GGAACAGCTCGCCCCCTCTGGAGGGGGCTCTCAACCTCTTTGA  
 359 N S S A P S G G G A S P N L S L T  
 1123 CTGAGAACATTCTGGAAACTACTCTGTGAGGCCAACACGGCTGGGG  
 376 A E H S G N Y S C E A A N G L G A  
 1174 CCCAGTGCAGTGAGGCAGTGCAGTCCATCTCAGGACCTGATGGCTATA  
 393 Q C S E A V P V S I S G P D G Y R  
 1225 GAAGAGACCTCATGACAGCTGGAGTTCTGGGGACTGTTGGTGTCTTG  
 410 R D L M T A G V L W G L F G V L G  
 1276 GTTCACTGGTGTGTTGTTGTTGATGCTTGTCCACAAGATATCAG  
 427 F T G V A L L Y A L F H K I S G  
 1327 GAGAAAGTCTGCCACTAATGAACCCAGAGGGCTCCAGGCCAACCTC  
 444 E S S A T N E P R G A S R P N P Q  
 1378 AAGAGTTACCTATTCAAGCCCAACCCAGACATGGAGGAGCTGCAGCCAG  
 461 E F T Y S S P T P D M E E L Q P V  
 1429 TGTATGTCAATGTGGCTCTGTAGATGTGGATGTGGTTATCTCAGGTCT  
 478 Y V N V G S V D V D V V Y S Q V W  
 1480 GGAGCATGCAGCAGCCAGAACAGTCAGCAAACATCAGGACACTCTGGAGA  
 495 S M Q Q P E S S A N I R T L L E N  
 1531 ACAAGGACTCCAAGTCATCTACTCTCTGTGAAGAAATCATAACACTTGG  
 512 K D S Q V I Y S S V K K S  
 1582 AGGAATCAGAAGGGAGATCAACAGCAAGGATGGGCATCATTAAGACTTG  
 1633 CTATAAAACCTTATGAAAATGTTGAGGCTTATCACCTGCCACAGCCAGAA  
 1684 CGTGCCTCAGGAGGCACCTCTGCTATTCTGCTCTGTGATGATGTTCTCT  
 1735 CCAATATCTCTTTACCTATCAATATTCTGACTGCTGCTACATCCAG  
 1786 AACTGTGCAAAATAAAATTCTGCTACCTCTCTTAAGGAAATCAGTGTG  
 1837 TAAAGATTGAGGGAGAAATGAATAAGAGATAACAGGTCTCACCTTATCT  
 1888 ACTGTGAAGTGATGAGAACAGGACTTGATAGTGTGATTAACCTTATTT  
 1939 GTGCTGCTGGATACAGTTGCTAATATTTGTTGAGAATTTCGAAATAT

32/34  
FIGURE 18D-2

1990 GTTCATTGGGAATATGGCTGAAATTTCTTCCACTGTGTCCTGCCA  
2041 GAATGTTGTATCAGGCTGATGCTGGCTCATAGAATGAGTTAGGCAGGAG  
2092 CCCTCCCTCTTGATTTTGGCATAGTTTCAGCAGGATGGTACCAAGTTA  
2143 TTCTTCTGCATCTGTAGAATTCACTATGAAATCCATCTGGTCTAGGGCT  
2194 TTGTGTTGGTTGGTAAGTTTTTATTACTAATTCAACTTCAGGCCCTGAT  
2245 ATTGGCTAGGACGGGTTCTGCTCTTCTGGTCAATCTGGAGATTC  
2296 TGCGTTCCAGGAATTAGCCGTTCTCCAGATTTCCTTTATGTGCA  
2347 TCGACTTGAGTGAAACATAACTTATATGCACGTGGAAACCAAAAATCTG  
2398 TGCGACTTGCTTTATTGCACGATTGGTTTATTGGTAGCTGGAACCTGA  
2449 ACCTGCAATATCACCAAAGTATGCATATAGTTGCAAAAATGTGATTTTGA  
2500 CATACTAAATATGAGTATTGCAATAAACTATGATATTACTTTGTAAGTA  
2551 TATAGAATAAAATGTAATAATCTATAAAA

33/34

## FIGURE 18E-1

1 GAGGCATCTAGGTACCATCCCTGACCTGGTCC  
 37 ATGCTGCCAGGGCTGTTGCTGTTGATCTGTGCTCCACTCTGTGAA  
 M L P R L L I C A P L C E  
 82 CCTGCCAGCTGTTTGATAGCAGCCCCCTCCATCCCCACAGAG  
 P A E L F L I A S P S H P T E  
 127 GGGAGCCCAGTGACCCCTGACGTGAAGATGCCCTTCTACAGAGT  
 G S P V T L T C K M P F L Q S  
 172 TCAGATGCCAGTTCCAGTTCTGCTTTCAGAGACACCCGGGCC  
 S D A Q F Q F C F R D T R A  
 217 TTGGGCCAGGCTGGAGCAGCTCCCCAAGCTCCAGATCGCTGCC  
 L G P G W S S S P K L Q I A A  
 262 ATGTTGAAAGAACAGACACAGGGTCAACTGGTGGCAGGGCACAGACA  
 M W K E D T G S Y W C E A Q T  
 307 ATGGCGTCCAAAGTCTGAGGAGCAGGAGATCCAGATAATGTG  
 M A S K V L R S R R S Q I N V  
 352 CACAGGGTCCCTGCGCTGATGTGAGCTGGAGACTCAGCCCCCA  
 H R V P V A D V S L E T Q P P  
 397 GGAGGACAGGTATGGAGGGAGACAGGGCTGGTCCCTCATGCTCA  
 G G Q V M E G D R L V L I C S  
 442 GTTGCTATGGCACAGGAGACATCACCTTCCCTGGTACAAAGGG  
 V A M G T G D I T F L W Y K G  
 487 GCTGTAGCTTAAACCTTCAGTCAGCAAAGACCCAGGGTCACTGACA  
 A V G L N L Q S K T Q R S L T  
 532 CCAGAGTATGAGATTCTTCAGTGAGGGAGAGTGTGAGCAA  
 A E Y E I P S V R E S D A E Q  
 577 TATTACTGTAGCTGAAATGCTATGTCAGCCCCAGTGG  
 Y Y C V A E N G Y G P S P S G  
 622 CTGGTGAGCATCACTGTCAGAAATCCCGGTGTCGCCCCAATCTC  
 L V S I T V R I P V S R P I L  
 667 ATGCTCAGGGCTCCAGGGCCAGGGCTGCAGTGGAGGATGTGCTG  
 M L R A P R A Q A A V E D V L  
 712 GAGCTTCACTGTGAGGCCCTGAGAGGCTCTCCCTCAATCTGTAC  
 E L H C E A L R G S P P I L Y  
 757 TGGTTTATCACGAGGATATCACCTGGGGAGCAGGTGGCCCC  
 W F Y H E D I T L G S R S A P  
 802 TCTGGAGGAGGAGCCTCCCTCACCTTCCCTGACTGAAGAACAT  
 S G G G A S F N L S L T E E H  
 847 TCTGGAAACTACTCCCTGTGAGGCCAACATGGCTGGGGCCAG  
 S G N Y S C E A N N G L G A Q  
 892 CGCAGTGAGGCGGTGACACTCAACTTCACAGTGCTACTGGGGCC  
 R S E A V T L N F T V P T G A  
 937 AGAAGCAATCATTTACCTTCAGGAGTCATTGAGGGCTGCTCAGC  
 R S N H L T S G V I E G L L S  
 982 ACCCTTGGTCCAGCCACGGTGGCTTATTTTGCTACGGCCTC  
 T L G P A T V A L L F C Y G L  
 1027 AAAAGAAAATAGGAAGACGTTCAAGCCAGGGATCCACTCAGGAGC  
 K R K I G R R S A R D P L R S  
 1072 CTCCCCAGCCCTCTACCCCAAGAGTTACCTCACCTCACCT  
 L P S P L P Q E F T Y L N S P  
 1117 ACCCCAGGGCAGCTACAGCCCTATATAGAAAATGTGAATGTTGTA  
 T P G Q L Q P I Y E N V N V V  
 1162 AGTGGGGATGAGGTATTACTGGCTACTATAACCAGCCGGAG  
 S G D E V Y S L A Y Y N Q P E  
 1207 CAGGAATCACTAGCAGCAGAAACCCCTGGGGACACATATGGAGGAC  
 Q E S V A A E T L G T H M E D  
 1252 AAGTTTCTTAGACATCTATCCAGGCTGAGGAAGCAAAACATT  
 K V S R L D I Y S R L R K A N I  
 1297 ACAGATGTGGACTATGAAGATGCTATGTA 1326  
 T D V D Y E D A M \*

GGTT ATGGAAGATT CTGCTTTG

1351 AAAACCATCC ATGACCCCAA GCCTCAGGCC TGATATGTT TTCAGAGATC  
 1401 CTGGGGCATT AGCTTCCAG TATACCTCTT CTGGATGCCA TTCTCCATGG  
 1451 CACTATTCTC TCATCTACTG TGAAGTGAAG TTGGGGCAGC CCTGAGAAA  
 1501 CTACCTAGGA GAACTAATAG ACACAGGAGT GACAGGACT TTGTTATCAG  
 1551 AACCAAGATT CTGCCGGCTC CTTGAAAAC AGGTCACTATT GTGCTCTTCT  
 1601 GTTTACAAGA GGAAACAAGA TGGAAATAAA GAAATTGGGA TCTTGGGTTG  
 1651 GAGGGACAGT GAAGCTTAGA GCACATGAAC TCAAGGTTAG TGACTCTGCA  
 1701 GGACTTCACA GAGAGAGCTG TGCCCCATCAT TCAGTCAGG TGCTTCTCT  
 1751 GCCCAGACAG CACAGAACTC CAGCCCCGCT ACTTCATGG ATCATCGAGT  
 1801 TTCCACCTAA AATATGATTC TATTTATTTT GAGTCAGTGT TACCAAATTA

**34/34**

**FIGURE 18E-2**

1851 GAACTAAAAC AAAGTTACAT AAAAAGTTAT TGIGACTCCA CTTAATTATA  
1901 GTGACGTATT TTGTTATATA TAGGCCAACC TATACCACAT CCAAAATTAT  
1951 GTATCTATTA CAGCCCCTAG AAGCTTTATA AATACAGTGT GTCTTCCTTT  
2001 ATTCCAAAAA TTTTGAAAT CGTGGTAATA TGTTTGAAA CCTGTATCTT  
2051 AATTATTTT TTGTTAAATT GAGACAGGGT CTCACTCTGT CACTCAATCT  
2101 GGAATGCAGT GGCAAAATCT TGCCCTCACTG CAACGCCCTGC CTCTCAGGCT  
2151 CAAGCAAACC TCTCACCTCA GCCTGCTGAG TAGCTGGGAC TACAGGCACA  
2201 TGCCACCAAA CTGGCCATT TTTTGTCTTA CGTAGAGACA AGATTTCACCC  
2251 GTTTGCCCAGGCTGGTCTC AAACCTCCGG GCTCAAGCAA TGTATTGAAT  
2301 TTT